82698

From:	Ibrahim, Medina A.
Sent:	Wednesday, December 18, 2002 6:49 PM
To:	STIC-Biotech/ChemLib
Subject:	09/538, 396

RECEIVED

DEC 19 2002

TECH/CHEM. DIVISION
(STIC)

Please search the following:

1. oligo search of at least 30 contiguous bases of SEQ ID NO:1.
2. residue of at least 20 contiguous amino acids of SEQ ID NO:2. Thanks

Medina A. Ibrahim
Patent Examiner
GAU:1638
CM1-9A12
(703)306-5822

Initial Contact:
Toby Post
Technical Info. Specialist
CM1 6A04
703-308-3534

	TYPE OF SEARCH:	VENDOR/COST (where applic.)
Searcher:_____	NA Sequences:_____	STN:_____
Phone:_____	AA Sequences:_____	DIALOG:_____
Location:_____	Structures:_____	Questel/Orbit:_____
Date Picked Up:___npo___	Bibliographic:_____	DRLink:_____
Date Completed:___npu___	Litigation:_____	Lexis/Nexis:_____
Searcher Prep/Review:_____	Full text:_____	Sequence Sys.:_____
Clerical:_____	Patent Family:_____	WWW/Internet:_____
Online time:_____	Other:_____	Other (specify):_____

OM nucleic - nucleic search, using sw model

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

Run on: December 26, 2002, 05:24:41 ; Search time 125 seconds
                                             (without alignments)
                                             14598.440 Million cell updates/sec

Title:
Perfect score: 4492
Sequence: US-09-538-396-1
                 1 aattcggcacgagtggatcc...........atctagcatcaaaaaaaaa 4492

Scoring table: OLIGO_NUC
                 Gapop 60.0 , Gapext 60.0

Searched: 363474 seqs, 20311208 residues .

Word size :   30

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :   Published_Applications_NA:*
                 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
                 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
                 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
                 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
                 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
                 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
                 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
                 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
                 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
                10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
                11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
                12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
                13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
                14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES

Result   &
 No.   Score   Query   DB   ID
         Match Length   Description

No matches found

Search completed: December 26, 2002, 08:43:45
Job time : 127 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 26, 2002, 08:43:51 ; Search time 27 Seconds
                                               (without alignments)
                                               1434.095 Million cell updates/sec

Title: US-09-538-396-2
Perfect score: 1316
Sequence: 1 MSTVDKMLIKGIRSFDPDNK..........YRVSKDENQHSIESQEIFD 1316

Scoring table: OLIGO
                Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size :   20

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :   Issued_Patents_AA:*
                1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
                2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
                3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
                4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
                5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
                6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES

Result   Query
No.   Score   Match Length DB   ID   Description
------ -----   ----- ------ --   --   -----------
1   31   2.4   1312   2   US-08-592-126-148   Sequence 148, App
2   31   2.4   1312   2   US-08-687-080-51   Sequence 51, Appl

                             ALIGNMENTS

RESULT 1
US-08-592-126-148
Sequence 148, Application US/08592126
Patent No. 5821091
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 151
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Dehlinger & Associates
;   STREET: 350 Cambridge Avenue, Suite 250
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94306
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;
;
;
;
;
;
;
;
;
;
;

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,126
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0960
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 148:
SEQUENCE CHARACTERISTICS:
LENGTH: 1312 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Rad50.pro-translation of SEQ ID NO:54
US-08-592-126-148

Qy   1207 MRGRCSAGQKVLASLIIRLALAETFCLNCGI 1237
            |||||||||||||||||||||||||||||||
Db   1197 MRGRCSAGQKVLASLIIRLALAETFCLNCGI 1227

Query Match   2.4%;   Score 31; DB 2;   Length 1312;
Best Local Similarity 100.0%; Pred. No. 4.8e-20;
Matches   31; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

RESULT 2
US-08-687-080-51
Sequence 51, Application US/08687080
Patent No. 5965427
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Dehlinger & Associates
;   STREET: 350 Cambridge Avenue, Suite 250
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94306
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,080
; FILING DATE: 17-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/592,126
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1312 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 26, 2002, 08:45:21 ; Search time 39 seconds
                                                 (without alignments)
                                                 584.241 Million cell updates/sec

Title:
Perfect score: 1316
Sequence: 1 MSTVDKMLIKGIRSFPDPDNK..........YRVSKDENQHSIIESQEIFD 1316

Scoring table: OLIGO
                  Gapop 60.0 , Gapext 60.0

Searched: 109717 seqs, 17314136 residues

Word size :   20

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :   Published_Applications_AA:*
                  1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
                  2: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
                  3: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
                  4: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
                  5: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
                  6: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
                  7: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
                  8: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
                  9: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
                 10: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
                 11: /cgn2_6/ptodata/1/pubpaa/US02_PUBCOMB.pep:*
                 12: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
                 13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
                 14: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
                     /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES

Result   %   Query
No.   Score   Match Length DB   ID   Description
------ -----   ------------ --   --   -----------

No matches found

Search completed: December 26, 2002, 08:52:12
Job time : 39 secs

OM protein - protein search, using sw model

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

Run on: December 26, 2002, 08:41:21 ; Search time 46 Seconds
                                              (without alignments)
                                              2750.280 Million cell updates/sec

Title: US-09-538-396-2
Perfect score: 1316
Sequence: 1 MSTVDKMLIKGIRSFDPDNK.........YRVSKDENQHSIIESQEIFD 1316

Scoring table: OLIGO
                 Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size :   20

Total number of hits satisfying chosen parameters: 2

Database :   PIR_73:*
                 1:   pir1:*
                 2:   pir2:*
                 3:   pir3:*
                 4:   pir4:*

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	% DB	ID	Description
1	49	3.7	1292	2	D84727	probable RAD50 DNA
2	42	3.2	1312	2	T30845	probable DNA repai

ALIGNMENTS

RESULT 1
D84727
probable RAD50 DNA repair protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 01-Mar-2002
C;Accession: D84727
R;Lin, X.; Kaul, S.; Rounsley, S. D.; Shea, T. P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: D84727
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1292 <STO>
A;Cross-references: GB:AE002093; NID:g4263721; PIDN:AAD15407.1; GSPDB:GN00139
A;Genetics:
A;Gene: At2g31970
A;Map position: 2
C;Superfamily: RAD50 protein

Qy   152 LENVIFVHQDESNWPLQDPSTLKKKFDDIFSATRYTKALEVIKKLHKDQ 200
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   152 LENVIFVHQDESNWPLQDPSTLKKKFDDIFSATRYTKALEVIKKLHKDQ 200

RESULT 2
T30845
probable DNA repair protein RAD50 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Mar-2000
C;Accession: T30845
R;Kim, K.K.; Daud, A.I.; Wong, S.C.; Pajak, L.; Tsai, S.C.; Wang, H.; Henzel, W.J.; F
J. Biol. Chem. 271, 29255-29264, 1996
A;Title: Mouse RAD50 has limited epitopic homology to p53 and is expressed in the adu
A;Reference number: Z20899; MUID:97067183; PMID:8910585
A;Accession: T30845
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1312 <KIM>
A;Cross-references: EMBL:U66887; NID:g1575574; PID:g1575575; PIDN:AAC52894.1
A;Genetics:
A;Gene: RAD50
A;Map position: 11
C;Superfamily: RAD50 protein
C;Keywords: DNA repair

Qy   1207 MRGRCSAGQKVLASLIIRLALAETFCLNGGILALDEPTTNLD 1248
          ||||||||||||||||||||||||||||||||||||||||||
Db   1197 MRGRCSAGQKVLASLIIRLALAETFCLNGGILALDEPTTNLD 1238

Search completed: December 26, 2002, 08:50:24
Job time : 46 secs

Query Match   3.7%;   Score 49;   DB 2;   Length 1292;
Best Local Similarity   100.0%;   Pred. No. 6.6e-40;
Matches   49;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Query Match   3.2%;   Score 42;   DB 2;   Length 1312;
Best Local Similarity   100.0%;   Pred. No. 7.2e-33;
Matches   42;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

OM protein - protein search, using sw model

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

Run on: December 26, 2002, 07:47:17 ; Search time 25 Seconds
                                                   (without alignments)
                                                   2183.314 Million cell updates/sec

Title:
Perfect score: 1316
Sequence: 1 MSTVDKMLIKGIRSFPDPDNK...........YRVSKDENQHSIIESQEIFD 1316

Scoring table: OLIGO
                     Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size :   20

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :   SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES

Result   %   Query
No.   Score Match Length DB   ID   Description
-----   ----- ----- ------ --   --   -----------

No matches found

Search completed: December 26, 2002, 08:47:34
Job time : 25 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2002, 05:15:21 ; Search time 104 Seconds
                                            (without alignments)
                                            13246.077 Million cell updates/sec

Title:
Perfect score: 4492
Sequence: 1 aattcggcacgagtggatcc.............atctagcatcaaaaaaaaaa 4492

Scoring table: OLIGO_NUC
                Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 15338381 residues

Word size :   30

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :   Issued_Patents_NA:*
                1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
                2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
                3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.

                                    SUMMARIES

Result   Query
No.   Score   Match Length DB ID   Description
------ -----   ------------ -- --   -----------

No matches found

Search completed: December 26, 2002, 08:41:16
Job time : 105 secs

OM protein - protein search, using sw model

```
                    GenCore version 5.1.3
             Copyright (c) 1993 - 2002 Compugen Ltd.

Run on:          December 26, 2002; 08:43:51 ; Search time 27 Seconds
                                     (without alignments)
                                     1434.095 Million cell updates/sec

Title:           US-09-538-396-2
Perfect score:   1316
Sequence:        1 MSTVDKMLIKGIRSFPDPDNK.........YRVSKDENQHSIIESQEIFD 1316

Scoring table:   OLIGO
                 Gapop 60.0 , Gapext 60.0

Searched:        262574 seqs, 29422922 residues

Word size :      20

Total number of hits satisfying chosen parameters:  2

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :       Issued_Patents_AA:*
                 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
                 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
                 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
                 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
                 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
                 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

```
Result                Query
No.     Score  Match Length  DB  ID               Description
------  -----  ------------  --  --------------   -----------
1       31     2.4    1312   2   US-08-592-126-148  Sequence 148, App
2       31     2.4    1312   2   US-08-687-080-51   Sequence 51, Appl
```

ALIGNMENTS

```
RESULT 1
US-08-592-126-148
Sequence 148, Application US/08592126
Patent No. 5821001
;  GENERAL INFORMATION:
;  APPLICANT: Gregory Dolganov
;  TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
;  TITLE OF INVENTION: Polypeptides
;  NUMBER OF SEQUENCES: 151
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE: Dehlinger & Associates
;  STREET: 350 Cambridge Avenue, Suite 250
;  CITY: Palo Alto
;  STATE: CA
;  COUNTRY: USA
;  ZIP: 94306
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE: Floppy disk
;  COMPUTER: IBM PC compatible
;  OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
;  SOFTWARE: Patentin Release #1.0, Version #1.25
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER: US/08/592,126
;  FILING DATE:
;  CLASSIFICATION: 435
;  ATTORNEY/AGENT INFORMATION:
;  NAME: Sholtz, Charles K.
;  REGISTRATION NUMBER: 38,615
;  REFERENCE/DOCKET NUMBER: 4600-0111
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE: (415) 324-0960
;  TELEFAX: (415) 324-0880
;  INFORMATION FOR SEQ ID NO: 148:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH: 1312 amino acids
;  TYPE: amino acid
;  TOPOLOGY: linear
;  MOLECULE TYPE: protein
;  HYPOTHETICAL: NO
;  ORIGINAL SOURCE:
;  INDIVIDUAL ISOLATE: Rad50.pro-translation of SEQ ID NO:54
US-08-592-126-148

Query Match            2.4%; Score 31; DB 2; Length 1312;
Best Local Similarity  100.0%; Pred. No. 4.8e-20;
Matches   31; Conservative  0; Mismatches  0; Indels  0; Gaps  0;

Qy   1207 MRGRCSAGQKVLASLIIRLALAETFCLNCGI 1237
          |||||||||||||||||||||||||||||||
Db   1197 MRGRCSAGQKVLASLIIRLALAETFCLNCGI 1227
```

```
RESULT 2
US-08-687-080-51
Sequence 51, Application US/08687080
Patent No. 5965427
;  GENERAL INFORMATION:
;  APPLICANT: Gregory Dolganov
;  TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
;  NUMBER OF SEQUENCES: 175
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE: Dehlinger & Associates
;  STREET: 350 Cambridge Avenue, Suite 250
;  CITY: Palo Alto
;  STATE: CA
;  COUNTRY: USA
;  ZIP: 94306
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE: Floppy disk
;  COMPUTER: IBM PC compatible
;  OPERATING SYSTEM: PC-DOS/MS-DOS
;  SOFTWARE: Patentin Release #1.0, Version #1.25
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER: US/08/687,080
;  FILING DATE: 17-JUL-1996
;  CLASSIFICATION: 435
;  PRIOR APPLICATION DATA:
;  APPLICATION NUMBER: US 08/592,126
;  FILING DATE: 26-JAN-1996
;  ATTORNEY/AGENT INFORMATION:
;  NAME: Sholtz, Charles K.
;  REGISTRATION NUMBER: 38,615
;  REFERENCE/DOCKET NUMBER: 4600-0111.30
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE: (415) 324-0880
;  TELEFAX: (415) 324-0960
;  INFORMATION FOR SEQ ID NO: 51:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH: 1312 amino acids
;  TYPE: amino acid
;  STRANDEDNESS: single
;  TOPOLOGY: linear
;  MOLECULE TYPE: protein
```

```
;    HYPOTHETICAL:      NO
;    ANTI-SENSE:        NO
;    ORIGINAL SOURCE:
;    INDIVIDUAL ISOLATE:    TRANS. OF RAD50 CDNA (SEQ. 54), NT.
;    INDIVIDUAL ISOLATE:    389 TO 4324
US-08-687-080-51

     Query Match          2.4%;    Score 31;   DB 2;   Length 1312;
     Best Local Similarity  100.0%;   Pred. No. 4.8e-20;
     Matches    31;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1207  MRGRCSAGQKVLASLIIRLALAEFFCLNCGI  237
                |||||||||||||||||||||||||||||||
Db   1197  MRGRCSAGQKVLASLIIRLALAEFFCLNCGI  1227


Search completed: December 26, 2002, 08:51:12
Job time : 27 secs
```

```
OM nucleic - nucleic search, using sw model

Run on:         December 26, 2002, 04:03:56 ; Search time 580 Seconds
                                              (without alignments)
                                              17441.339 Million cell updates/sec

                GenCore version 5.1.3
                Copyright (c) 1993 - 2002 Compugen Ltd.

Title:          us-09-538-396-1
Perfect score:  4492
Sequence:       1 aattcggcacgagtggatcc...........atctagcatcaaaaaaaaaa 4492

Scoring table:  OLIGO_NUC
                Gapop 60.0 , Gapext 60.0

Searched:       2185239 seqs, 112599159 residues

Word size :     30

Total number of hits satisfying chosen parameters:      1

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :    N_Geneseq_101002:*
          1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
          2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
          3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
          4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
          5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
          6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
          7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
          8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
          9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
         10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
         11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
         12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
         13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
         14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
         15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
         16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
         17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
         18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
         19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
         20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
         21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
         22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
         23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
         24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
Result                   %
No.    Score  Match  Query    DB    ID
              Length DB
-----  -----  ------ -----   ----  --------          Description
  1    4492   100.0  4492     22   AAC66915          Maize RAD50 coding
```

ALIGNMENTS

```
RESULT 1
AAC66915
ID   AAC66915 standard; cDNA: 4492 BP.
```

```
XX
AC   AAC66915;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Maize RAD50 coding sequence.
XX
KW   Maize; RAD50; meiotic recombination; DNA repair; chromosome 4p;
KW   recombination efficiency; transformation efficiency; ss.
XX
OS   zea mays.
XX
PN   WO200068404-A1.
XX
PD   16-NOV-2000.
XX
PF   25-APR-2000; 2000WO-US11086.
XX
PR   05-MAY-1999; 99US-0132575.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Mahajan PB, Shi J;
XX
DR   WPI; 2001-007402/01.
DR   P-PSDB; AAB27248.
XX
PT   Nucleic acids encoding Rad50 polypeptides, useful for transgenic plant
PT   production and for producing antibodies that are useful for modifying
PT   gene expression -
XX
PS   Claim 1; Page 65-70; 76pp; English.
XX
CC   The present sequence is the coding sequence of the maize RAD50 gene,
CC   found on maize chromosome 4p. RAD50 is involved in meiotic recombination
CC   and DNA repair, and the gene and its protein can be used to produce
CC   transgenic plants whose expression of the gene may be regulated. This can
CC   be useful in the regulation of transformation and recombination
CC   efficiency in plants.
XX
SQ   Sequence 4492 BP; 1537 A; 831 C; 1061 G; 1063 T; 0 other;

Query Match             100.0%;  Score 4492;  DB 22;  Length 4492;
Best Local Similarity   100.0%;
Matches 4492;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 AATTCGGCACGAGTGGATCCAATTAGCACCCATAGCCGTACGAAAAACCCTAAGAACCCTAAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AATTCGGCACGAGTGGATCCATTAGCACCCATAGCCGTACGAAAAACCCTAAGAACCCTAAC 60

Qy         61 CGGTACAAAAACCCTAAAAACCCTAACGGCGCGGCTGAGAACTCCAAAAAAACGCGATTTCTC 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGTACAAAAACCCTAAAAACCCTAACGCCGCTGAGAACTCCAAAAAAACGCGATTTCTC 120

Qy        121 CTCCACTGCGCCCCTCCTTTTCTCTTTCCAATGTGTTTTGCAATCACTACGACGGTAATGAAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCCACTGCGCCCCTCCTTTTCTCTTTCCAATGTGTTTTGCAATCACTACGACGGTAATGAAT 180

Qy        181 AGAAGTTGATAGGGAGGAGGAGGAGCACTTGGGCTCGTGGGGCTCGTGGGCCAGAC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGAAGTTGATAGGGAGGAGGAGGAGCACTTGGGCTCGTGGGGCTCGTGGGCCAGAC 240

Qy        241 TGGATCGGAGTGCAAGTCGTAGAGGAGGGCACTTGGGGCTCGTGGGGCTCGTGGGCCAGAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGATCGGAGTGCAAGTCGTAGAGGAGGGCACTTGGGGCTCGTGGGGCTCGTGGGCCAGAC 300

Qy        301 GTTGACAAAGATGCTGATCAAGGGGATTCGGAGACTTCGATCCGGACAAATAAGAACGTCATC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTTGACAAAGATGCTGATCAAGGGGATTCGGAGACTTCGATCCGGACAAATAAGAACGTCATC 360

Qy        361 ACCTTCTTCAAGCCGCTCACCCTCATCGTTGGCCCCAACGCGTGCTGCCAAGACCACGATC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACCTTCTTCAAGCCGCTCACCCTCATCGTTGGCCCCAACGCGTGCTGCCAAGACCACGATC 420
```

```
Qy    421 ATCGAGTGCCTGAAGCTTTCTTGCACCGGCGAGCTGCCCCCCAACTCCCGCTCTGGCCAC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATCGAGTGCCTGAAGCTTTCTTGCACCGGCGAGCTGCCCCCCAACTCCCGCTCTGGCCAC 480

Qy    481 ACCTTCGTCCACGACCCCAAGGTAGCTGGCGAGACGGAAACAAAAGGACAAATTAAGTTG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACCTTCGTCCACGACCCCAAGGTAGCTGGCGAGACGGAAACAAAAGGACAAATTAAGTTG 540

Qy    541 CGGTTTAAGACTGCAGCAGGAAAGGATGTGGTGTGCATCCGGTCCTTCCAGCTTACCCAA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CGGTTTAAGACTGCAGCAGGAAAGGATGTGGTGTGCATCCGGTCCTTCCAGCTTACCCAA 600

Qy    601 AAGGCATCAAAGATGGAGTTTAAGGCAATTGAAAGCGTCCTCCAGACTATAAATCCACAC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AAGGCATCAAAGATGGAGTTTAAGGCAATTGAAAGCGTCCTCCAGACTATAAATCCACAC 660

Qy    661 ACAGGGGAGAAAGTCTGCCTCAGCTACAGATGTGCTGACATGGATAGAGAGATTCCTGCC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ACAGGGGAGAAAGTCTGCCTCAGCTACAGATGTGCTGACATGGATAGAGAGATTCCTGCC 720

Qy    721 TTAATGGGTGTTTCGAAGGCCGTACTGGAGAATGTTATATTTGTTCACCAAGATGAATCC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TTAATGGGTGTTTCGAAGGCCGTACTGGAGAATGTTATATTTGTTCACCAAGATGAATCC 780

Qy    781 AATTGGCCATTGCAGGACCCGTCAACACTTAAGAAGAAGTTCGATGACATCTTCTCTGCC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AATTGGCCATTGCAGGACCCGTCAACACTTAAGAAGAAGTTCGATGACATCTTCTCTGCC 840

Qy    841 ACACGCTATACGAAAGCTCTTGAAGTCATAAAGAAACTTCACAAGGATCAAATGCAAGAG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ACACGCTATACGAAAGCTCTTGAAGTCATAAAGAAACTTCACAAGGATCAAATGCAAGAG 900

Qy    901 ATCAAGACTTTTAGGTTAAAGCTGGAGAACCTTCAGACTGTAAAAGACCAAGCACATAAG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ATCAAGACTTTTAGGTTAAAGCTGGAGAACCTTCAGACTGTAAAAGACCAAGCACATAAG 960

Qy    961 CTGCGTGAAAATATTGCTCAAGATCAAGAAAAGTCAGATGCCTCAAAATCTCAGATGGAG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CTGCGTGAAAATATTGCTCAAGATCAAGAAAAGTCAGATGCCTCAAAATCTCAGATGGAG 1020

Qy   1021 CAACTGAAGGAAAAGATCTGTGGTACCGAGAGAGAAATCCTGCAAATGGAAACAAGTTTG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CAACTGAAGGAAAAGATCTGTGGTACCGAGAGAGAAATCCTGCAAATGGAAACAAGTTTG 1080

Qy   1081 GATGAACTGAGAAGACTTCAGGGACAAATTGACATCAAGGCAACAGAGAGAAGTACATTA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GATGAACTGAGAAGACTTCAGGGACAAATTGACATCAAGGCAACAGAGAGAAGTACATTA 1140

Qy   1141 CTTACGCAGCAGCATGAAAAGCTTGCTGCACTTTCTGAGGAAAATGAAGATACCGATGAG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CTTACGCAGCAGCATGAAAAGCTTGCTGCACTTTCTGAGGAAAATGAAGATACCGATGAG 1200

Qy   1201 GAACTAATGGAATGGCAAACAAAATTTGAAGAAAGGATTGCGTTACTAGAAACAAAAATC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GAACTAATGGAATGGCAAACAAAATTTGAAGAAAGGATTGCGTTACTAGAAACAAAAATC 1260

Qy   1261 AGTAAACTTGTAAGAGATATGGATGATGAAGCATCTTATAGCTCCGTTCTGTCCAAACAA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AGTAAACTTGTAAGAGATATGGATGATGAAGCATCTTATAGCTCCGTTCTGTCCAAACAA 1320

Qy   1321 AATTCTGAATTAACACATGAAATTGGAAAGCTCCAGGCAGAAGCTGATGCTCACCTGACT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AATTCTGAATTAACACATGAAATTGGAAAGCTCCAGGCAGAAGCTGATGCTCACCTGACT 1380

Qy   1381 ATGAAGCATGAACGAGACTCAGACATAAAAAATATATGCACTAAACATAATCTTGGGCCG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 ATGAAGCATGAACGAGACTCAGACATAAAAAATATATGCACTAAACATAATCTTGGGCCG 1440

Qy   1441 GTTCCTGAACATCCCTTTACGAATGATGTTGCTATGAACCTTACAAACAGGATTAAAGCG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GTTCCTGAACATCCCTTTACGAATGATGTTGCTATGAACCTTACAAACAGGATTAAAGCG 1500

Qy   1501 AGACTATCAAGTCTTGAGAATGATTTGCTGGATAAGAAGAAATCCAATGAAGATCAGTTA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 AGACTATCAAGTCTTGAGAATGATTTGCTGGATAAGAAGAAATCCAATGAAGATCAGTTA 1560

Qy   1561 GATGTTTTGTGGAAACACTATCTTAAAATAAATGCTCGCTACTCCGAAGTTGATGGTCAG 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GATGTTTTGTGGAAACACTATCTTAAAATAAATGCTCGCTACTCCGAAGTTGATGGTCAG 1620

Qy   1621 ATACAATCTAAGATTGAATCCATGTCAGGCATTTTAAGACGGAGAAAAGATAAAGAGAAA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 ATACAATCTAAGATTGAATCCATGTCAGGCATTTTAAGACGGAGAAAAGATAAAGAGAAA 1680

Qy   1681 GAACGCGATGCTGCAGAAGTGGAGCTTTCAAAATTTAATCTATCCCGTATCGATGAGAGG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GAACGCGATGCTGCAGAAGTGGAGCTTTCAAAATTTAATCTATCCCGTATCGATGAGAGG 1740

Qy   1741 GAGAGACATATGCAAATTGAAGTCGAGAGGAAGACACTTGCGCTTGGAGAAAGAGACTAT 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GAGAGACATATGCAAATTGAAGTCGAGAGGAAGACACTTGCGCTTGGAGAAAGAGACTAT 1800

Qy   1801 GATTCAATTATAAGTCAGAAACGAACAGAAGTATATAGTTTGGAACAGAAAATAAAAGTG 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 GATTCAATTATAAGTCAGAAACGAACAGAAGTATATAGTTTGGAACAGAAAATAAAAGTG 1860

Qy   1861 CTTCTGCGGGAGAAAGATATAATAAATAGAAATGCTGATGAAAGAGTAAAACTGGGTTTG 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 CTTCTGCGGGAGAAAGATATAATAAATAGAAATGCTGATGAAAGAGTAAAACTGGGTTTG 1920

Qy   1921 AAGAAGGATGCATTGGAAAGCAGCAAGGACAAGCTCAATGAGATAGTTAATGAGCATAAG 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AAGAAGGATGCATTGGAAAGCAGCAAGGACAAGCTCAATGAGATAGTTAATGAGCATAAG 1980

Qy   1981 GATAAAATCAAAAAGGTACTTAGGGGGAGGAATCCTTTTGAGAAGGATATGAAGAAGGAG 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 GATAAAATCAAAAAGGTACTTAGGGGGAGGAATCCTTTTGAGAAGGATATGAAGAAGGAG 2040

Qy   2041 ATCAATCAAGCCTTTTGGCCTGTGGACAAGGAATACAATGAGTTAAGATCAAAATCCCAG 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 ATCAATCAAGCCTTTTGGCCTGTGGACAAGGAATACAATGAGTTAAGATCAAAATCCCAG 2100

Qy   2101 GAAGCAGAGCAAGAGCTTAAATTTACTCAGAGCAAAGTAACTGATGCTAGAGAACAATTG 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GAAGCAGAGCAAGAGCTTAAATTTACTCAGAGCAAAGTAACTGATGCTAGAGAACAATTG 2160

Qy   2161 ACAAAACTTCGAAGAGATATGGATGCAAAAAGAAGATTCCTGGACTCGAAACTTCAATCT 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ACAAAACTTCGAAGAGATATGGATGCAAAAAGAAGATTCCTGGACTCGAAACTTCAATCT 2220

Qy   2221 ATTTTACAGATATCTGCTAATGTTGACATGTTTCCCAAAGTTCTACAAGACGCCATGAAC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 ATTTTACAGATATCTGCTAATGTTGACATGTTTCCCAAAGTTCTACAAGACGCCATGAAC 2280

Qy   2281 AAAAGAGATGAACAGAAAAGATTAGAGAATTTCGCAAATGGAATGCGGGAAATGCTTGCA 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 AAAAGAGATGAACAGAAAAGATTAGAGAATTTCGCAAATGGAATGCGGGAAATGCTTGCA 2340

Qy   2341 CCTTTTGAACATTTGGCTCGGAAGAATCATGTATGCCCATGCTGTGAACGTGCTTTCACA 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 CCTTTTGAACATTTGGCTCGGAAGAATCATGTATGCCCATGCTGTGAACGTGCTTTCACA 2400

Qy   2401 CCTGATGAGGAGGATGAGTTCGTGAAGAAACAAAGGATGCAAAACTCAAGTACTGCAGAG 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 CCTGATGAGGAGGATGAGTTCGTGAAGAAACAAAGGATGCAAAACTCAAGTACTGCAGAG 2460

Qy   2461 AGATCTAAAGCTCTGGCAATGGAATCATCAAATGCTGAAGCTCTTTTTCAGCAATTGGAT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 AGATCTAAAGCTCTGGCAATGGAATCATCAAATGCTGAAGCTCTTTTTCAGCAATTGGAT 2520

Qy   2521 AAACTTCGGACTATCTATGATGCTTATGTGAAGCTGGTAGAAGAAACCATACCTCTAGCA 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 AAACTTCGGACTATCTATGATGCTTATGTGAAGCTGGTAGAAGAAACCATACCTCTAGCA 2580

Qy   2581 GAGAAAAACTTGAATCAACATTTGGCGGATGAAAGTCAGAAGGCGCAGGCATTTGATGAT 2640
```

```
Qy  2581 GAGAAAAACTTGAATCAACACATTTGGCGGATGAAAGTTCAGAAGGCGCAGGCATTTGATGAT 2640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641 CTTTTTGGGTGTGTTCTTGCCCATGTTCAAAATGACAGGGATGCAGTGGAAGCCTTATTACAA 2700
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2641 CTTTTTGGGTGTGTTCTTGCCCATGTTCAAAATGACAGGGATGCAGTGGAAGCCTTATTACAA 2700
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701 CCCACTGATACTATTGACACAGCCATGTACACATGTAAAATTCAACAGCTAGTCAAAGAAGTAGAA 2760
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2701 CCCACTGATACTATTGACACAGCCATGTACACATGTAAAATTCAACAGCTAGTCAAAGAAGTAGAA 2760
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761 GATCTTGAATATGCACTTGATTCTACTGTGCCCAGAAATTCAA 2820
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2761 GATCTTGAATATGCACTTGATTCTACTGTGCCCAGAAATTCAA 2820
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821 CTGGACCTGAACACTTTCTGACGAACAACAGGCACACATTGATTGGCCGAGTGGTCAAGTCTTTGGAGGAAATTCA 2880
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2821 CTGGACCTGAACACTTTCTGACGAACAACAGGCACACATTGATTGGCCGAGTGGTCAAGTCTTTGGAGGAAATTCA 2880
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881 AGAGATCAAACATAGAATGCTAAATGATCAAGTGCTCAGGTGCAGAGGCACAAT 2940
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2881 AGAGATCAAACATAGAATGCTAAATGATCAAGTGCTCAGGTGCAGAGGCACAAT 2940
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941 GCTCGGGAAGAGAAAGTGAAAGCTTCTAGCATATTGGAAAGAATTCCAAAAATCTGAAGAG 3000
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2941 GCTCGGGAAGAGAAAGTGAAAGCTTCTAGCATATTGGAAAGAATTCCAAAAATCTGAAGAG 3000
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001 GAATTGGTGCTTCTTCTGACGTGAGAAAAAAGAACAACTGATTGTAGAAGAAGCTTTAGAA 3060
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3001 GAATTGGTGCTTCTTCTGACGTGAGAAAAAAGAACAACTGATTGTAGAAGAAGCTTTAGAA 3060
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061 GAGTCTCTTGATCCATTGTCCAAAGAGAAAGAGAGCTTGTGTTGCAAGATAATGCTTTG 3120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3061 GAGTCTCTTGATCCATTGTCCAAAGAGAAAGAGAGCTTGTGTTGCAAGATAATGCTTTG 3120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121 AAGCAAAAGCTGGATGAAGAGAGTATCATCAGCTTGCACAGAAGAAAGAAAAAGGAGTTCCAGCAA 3180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3121 AAGCAAAAGCTGGATGAAGAGAGTATCATCAGCTTGCACAGAAGAAAGAAAAAGGAGTTCCAGCAA 3180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181 GAACTTGATGCTCTTGGAAGACTTAATATGAAGATAAAAAGGGTTACTTGGAATTCCAAGAAA 3240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3181 GAACTTGATGCTCTTGGAAGACTTAATATGAAGATAAAAAGGGTTACTTGGAATTCCAAGAAA 3240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241 AACGAAAAAGCTTAAGGAATTGCAGGGAAGGCATGTTCTTCTTTGCCCATTCTCAGTTACAGAGT 3300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3241 AACGAAAAAGCTTAAGGAATTGCAGGGAAGGCATGTTCTTCTTTGCCCATTCTCAGTTACAGAGT 3300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301 TGCCATGGCAAAACAGCAAAGAAATATCAGCTGAGTTAAACAAGAAGAACTACTGCAG 3360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3301 TGCCATGGCAAAACAGCAAAGAAATATCAGCTGAGTTAAACAAGAAGAACTACTGCAG 3360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361 GGCCAGGGCCAGTTGAAAAGAACATTGATGACAAATCTCAAGTACAGAAGAAACAAAGGCT 3420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3361 GGCCAGGGCCAGTTGAAAAGAACATTGATGACAAATCTCAAGTACAGAAGAAACAAAGGCT 3420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421 GARGTGGAACAACTTACTCGTGATATAGAATCACTTGAAGAACGGCTGAAATAGGT 3480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3421 GARGTGGAACAACTTACTCGTGATATAGAATCACTTGAAGAACGGCTGAAATAGGT 3480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481 AGCTTGTCTGCTATAGAGAGCTGATCTGAAACGCCATTCTCTCAAGAGAAGAGAGGCTTAAT 3540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3481 AGCTTGTCTGCTATAGAGAGCTGATCTGAAACGCCATTCTCTCAAGAGAAGAGAGGCTTAAT 3540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541 TCAGAATTTAACAGGTGGCAAGGAACACTTTCTGTTTATCAAAGTAATAATTTCAAAGCAC 3600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3541 TCAGAATTTAACAGGTGGCAAGGAACACTTTCTGTTTATCAAAGTAATAATTTCAAAGCAC 3600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601 AAACAAGAGACTTAAACTGTCACAGTACTACAAGGCGATATCGAGAAGAGAGTGTCAAATTT 3660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3601 AAACAAGAGACTTAAACTGTCACAGTACTACAAGGCGATATCGAGAAGAGAGTGTCAAATTT 3660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661 CTCCAGCTTAAGACAACTGAAAATGGCAAACAAGGACTTGGACAGAGATAATTATACTGCTTTA 3720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3661 CTCCAGCTTAAGACAACTGAAAATGGCAAACAAGGACTTGGACAGAGATAATTATACTGCTTTA 3720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721 GACAAGGCTCTTAATGCGTTCCACAGCATGAACAGAATGAGAGATGAAGATGAGGAGTAAAATAATAATCAAG 3780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3721 GACAAGGCTCTTAATGCGTTCCACAGCATGAACAGAATGAGAGATGAAGATGAGGAGTAAAATAATAATCAAG 3780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781 GAACTGTGGCAACAGACAATACAGAGGCCAGAGATATTGATTACAGATAAAATTCTGAT 3840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3781 GAACTGTGGCAACAGACAATACAGAGGCCAGAGATATTGATTACAGATAAAATTCTGAT 3840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841 TCTGAGGGTGGCTGGGCCACTCGATCATACAGCTACCGCGTGTTGTTATGCAAACTGTGTGATGCT 3900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3841 TCTGAGGGTGGCTGGGCCACTCGATCATACAGCTACCGCGTGTTGTTATGCAAACTGTGTGATGCT 3900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901 GAGCTGGGAAAATGCCAGGGGCGCTGCGAGTCTGCAGAGAGGTTCTTGCTTCTCTTATAATC 3960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3901 GAGCTGGGAAAATGCCAGGGGCGCTGCGAGTCTGCAGAGAGGTTCTTGCTTCTCTTATAATC 3960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961 AGACTAGCACTTGGGAAACTTTCTGCCTGAAACTGCTGGGTATAATCGAGTCCAAGGATGAG 4020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3961 AGACTAGCACTTGGGAAACTTTCTGCCTGAAACTGCTGGGTATAATCGAGTCCAAGGATGAG 4020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021 ACTACGAATCTAGAATGCAGAAAAGTCAGAGAGTCTTGCTCGGTGTTGAGAATAATG 4080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4021 ACTACGAATCTAGAATGCAGAAAAGTCAGAGAGTCTTGCTCGGTGTTGAGAATAATG 4080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081 GAAGCCAGGAGAAAAGGGCCAGGAGGAACTTCCAGTTGATTGGTGAATCACTCAIGATGAGAGATTT 4140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4081 GAAGCCAGGAGAAAAGGGCCAGGAGGAACTTCCAGTTGATTGGTGAATCACTCAIGATGAGAGATTT 4140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141 GCCCATCTTATCGGTCAAAGGCAGCTTGCTGAGAGTACTATTCGAGTCTCCAAGGATGAG 4200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4141 GCCCATCTTATCGGTCAAAGGCAGCTTGCTGAGAGTACTATTCGAGTCTCCAAGGATGAG 4200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201 AACCAGCACAGCATAATTGAATCCCAAGAAGATATTTGACTAAGGTGTTCTTAGGAGGTG 4260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4201 AACCAGCACAGCATAATTGAATCCCAAGAAGATATTTGACTAAGGTGTTCTTAGGAGGTG 4260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4261 TAGCACGCACTCGGTTCGTTGCTAGTCGAATCAGTTAATTAATTATGCCAAGTACTGGGTGCCAGAG 4320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4261 TAGCACGCACTCGGTTCGTTGCTAGTCGAATCAGTTAATTAATTATGCCAAGTACTGGGTGCCAGAG 4320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321 CAAATGTTAACAAGCTTTAACGGTCTCTGTTACGCCGTTAGTCAGTTGCACTGGCTGGAAACCTAT 4380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4321 CAAATGTTAACAAGCTTTAACGGTCTCTGTTACGCCGTTAGTCAGTTGCACTGGCTGGAAACCTAT 4380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381 CCTTCGTTGTTGTTATACTTATTTAATCTGGCCAGAGGATGGAAATGTGTGCACTGGGGTGAT 4440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4381 CCTTCGTTGTTGTTATACTTATTTAATCTGGCCAGAGGATGGAAATGTGTGCACTGGGGTGAT 4440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441 GGATGTTTCACGACATCAAATGAAATGTTTCACAATCTAGCATCAAAAAAAAAAAAA 4492
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4441 GGATGTTTCACGACATCAAATGAAATGTTTCACAATCTAGCATCAAAAAAAAAAAA 4492
```

OM nucleic - nucleic search, using sw model

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

Run on: December 26, 2002, 05:13:51 ; Search time 3869 seconds
            (without alignments)
            18803.341 Million cell updates/sec

Title: US-09-538-396-1
Perfect score: 4492
Sequence: 1 aattcggcacgagtggatcc..............atctagcatcaaaaaaaaaa 4492

Scoring table: OLIGO_NUC
                Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size :   30

Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :   EST:*
                1: em_estba:*
                2: em_esthum:*
                3: em_estin:*
                4: em_estmu:*
                5: em_estov:*
                6: em_estpl:*
                7: em_estro:*
                8: em_htc:*
                9: gb_est1:*
                10: gb_est2:*
                11: gb_htc:*
                12: gb_est3:*
                13: gb_est4:*
                14: gb_est5:*
                15: em_estfun:*
                16: em_estom:*
                17: gb_gss:*
                18: em_gss_hum:*
                19: em_gss_inv:*
                20: em_gss_pln:*
                21: em_gss_vrt:*
                22: em_gss_fun:*
                23: em_gss_mam:*
                24: em_gss_nus:*
                25: em_gss_other:*
                26: em_gss_pro:*
                27: em_gss_rod:*

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c						
1	131	2.9	549	17	BH776364	BH776364 fzmb013f0
2	99	2.2	512	12	BG649394	BG649394 EM1_78_H0
3	65	1.4	484	12	BG649231	BG649231 EM1_78_H0
4	47	1.0	416	13	BM330353	BM330353 PIC1_50_G
5	35	0.8	549	13	BI416347	BI416347 LjNEST2d3
6	30	0.7	458	10	BE443510	BE443510 WHE1115_D

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1
BH776364/c
LOCUS   BH776364   549 bp   DNA   linear   GSS 28-MAR-2002
DEFINITION fzmb013f00d607f0 fzmb filtered library Zea mays genomic clone
            fzmb013f00d607 5', DNA sequence.
ACCESSION   BH776364
VERSION   BH776364.1   GI:19778667
KEYWORDS   GSS.
SOURCE   Zea mays.
  ORGANISM Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 549)
  AUTHORS   Budiman,M.A., Freese,R.G., Bedell,J.A., Nunberg,A.N. and Lakey,N.D.
  TITLE   GeneThresher methylation filtered genomic sequences from maize
  JOURNAL   Unpublished (2002)
COMMENT   Contact: Bedell JA
            Orion Genomics, LLC
            4041 Forest Park Ave, St. Louis, MO 63108, USA
            Tel: 314 615 6979
            Fax: 314 615 5975
            Email: jbedell@oriongenomics.com
            Plate: fzmb013f006 row: d column: 07
            Seq primer: M13 forward
            Class: shotgun
            High quality sequence stop: 549.
FEATURES   Location/Qualifiers
     source   1..549
                     /organism="Zea mays"
                     /cultivar="Mo17"
                     /db_xref="taxon:4577"
                     /clone="fzmb013f006d07"
                     /clone_lib="fzmb filtered library"
                     /note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA
                     prepared from purified nuclei was randomly sheared,
                     end-repaired, size fractionated to enrich for the 0.5 to
                     5 kb fraction, ligated into HincII-digested pBCSK(-)
                     vector and electroporated into E. coli cells."
BASE COUNT   154 a   131 c   110 g   154 t
ORIGIN

Query Match   2.9%; Score 131; DB 17; Length 549;
Best Local Similarity   100.0%; Pred. No. 1.3e-52;
Matches 131; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   4072 AGAATAATGGAAGCCAGGAAAAGGGCAGGAGAAACTTCCAGTTGATTGTAATCACTCATGAT 4131
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   350 AGAATAATGGAAGCCAGGAAAAGGGCAGGAGAAACTTCCAGTTGATTGTAATCACTCATGAT 291

Qy   4132 GAGAGATTTGCCCATCTTATCGGTCAAAGGCAGCCTTGCTGAGAAGTACTATCGAGTCTCC 4191
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   290 GAGAGATTTGCCCATCTTATCGGTCAAAGGCAGCCTTGCTGAGAAGTACTATCGAGTCTCC 231

RESULT 2
BG649394
LOCUS   BG649394   512 bp   mRNA   linear   EST 24-APR-2001
DEFINITION EM1_78_H06.b1_A002 Embryo 1 (EM1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   BG649394
VERSION   BG649394.1   GI:13784506
KEYWORDS   EST.
SOURCE   sorghum.

Qy   4192 AAGGATGAGAA 4202
          |||||||||||
Db   230 AAGGATGAGAA 220

ORGANISM   Sorghum bicolor
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
           clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 512)
  AUTHORS Reid,S.P., Cordonnier-Pratt,M.-M., Gingle,A. and Pratt,L.H.
  TITLE   An EST database from Sorghum: developing embryos
  JOURNAL Unpublished (2000)
  COMMENT Contact: Cordonnier-Pratt MM
           Laboratory for Genomics and Bioinformatics
           The University of Georgia, Department of Plant Biology
           Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
           Tel: 706 542 1860
           Fax: 706 583 0210
           Email: mmpratt@uga.edu
           Sequences have been trimmed to exclude PolyA, vector and regions
           below Phred quality 16. The threshold for highest quality sequence
           is 20.
           Seq primer: JEN REV
           High quality sequence stop: 383
           POLYA=No.

FEATURES   Location/Qualifiers
     source   1..512
                     /organism="Sorghum bicolor"
                     /db_xref="taxon:4558"
                     /clone_lib="Embryo 1 (EM1)"
                     /note="Organ: Embryos germinated for 24 hr; Vector:
                     pBluescript II from Lambda Zap II; Site_1: XhoI; Site_2:
                     EcoRI; The library was made from poly-A RNA in the cloning
                     vector lambda ZAP II. Clones to be sequenced were
                     prepared by mass excision."
BASE COUNT   102 a   168 c   157 g   85 t
ORIGIN

Query Match   2.2%;   Score 99;   DB 12; Length 512;
Best Local Similarity   99.3%;   Pred. No. 6.5e-37;
Matches 149;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

Qy   3726 GGCTCTTATGCGGTTCCACAGCATGAAGATGGAGGAGGAGATAAATAAAATAATCAAGGAACT 3785
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   358 GGCTCTTATGCGGTTCCACAGCATGAAGATGGAGGAGGAGATAAATAAAATAATCAAGGAACT 417

Qy   3786 GTGGCAACAGACAATACAGAGGCCAGATATTGATTACACATAAGGCATAAAATCTGATTCTGA 3786
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   418 GTGGCAACAGACAATACAGAGGCCAGATATTGATTACACATAAGGCATAAAATCTGATTCTGA 418

Qy   3846 GGGTGCTGCTGGCACTCGATCATACAGCTACCG 3875
           |||||||||||||||||||||||||||||||||
Db   478 GGGTGCTGGCACTCGATCATACAGCTACCG 507

RESULT 3
LOCUS   BG649231   484 bp   mRNA   linear   EST 24-APR-2001
DEFINITION EM1_78_H06_g1_A002 Embryo 1 (EM1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   BG649231
VERSION   BG649231.1   GI:13784343
KEYWORDS   EST.
SOURCE   sorghum.
  ORGANISM Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1 (bases 1 to 484)
  AUTHORS   Reid,S.P., Cordonnier-Pratt,M.-M., Gingle,A. and Pratt,L.H.
  TITLE   An EST database from Sorghum: developing embryos
  JOURNAL   Unpublished (2000)
  COMMENT   Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector, and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: T7
            High quality sequence stop: 308
            POLYA=Yes.

FEATURES   Location/Qualifiers
     source   1..416
                     /organism="Sorghum bicolor"
                     /cultivar="BTx623"
                     /db_xref="taxon:4558"

REFERENCE 1 (bases 1 to 484)
  AUTHORS Reid,S.P., Cordonnier-Pratt,M.-M., Gingle,A. and Pratt,L.H.
  TITLE   An EST database from Sorghum: developing embryos
  JOURNAL Unpublished (2000)
  COMMENT Contact: Cordonnier-Pratt MM
           Laboratory for Genomics and Bioinformatics
           The University of Georgia, Department of Plant Biology
           Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
           Tel: 706 542 1860
           Fax: 706 583 0210
           Email: mmpratt@uga.edu
           High quality sequence start: 7
           High quality sequence stop: 484
           POLYA=No.

FEATURES   Location/Qualifiers
     source   1..484
                     /organism="Sorghum bicolor"
                     /db_xref="taxon:4558"
                     /clone_lib="Embryo 1 (EM1)"
                     /note="Organ: Embryos germinated for 24 hr; Vector:
                     pBluescript II from Lambda Zap II; Site_1: XhoI; Site_2:
                     EcoRI; The library was made from poly-A RNA in the cloning
                     vector lambda ZAP II. Clones to be sequenced were
                     prepared by mass excision."
BASE COUNT   145 a   92 c   113 g   134 t
ORIGIN

Query Match   1.4%;   Score 65;   DB 12; Length 484;
Best Local Similarity   100.0%;   Pred. No. 3.1e-20;
Matches   65;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   4066 CTGTTGAGAATAATGGAAGCCAGGAAAGGGCAGAGAGAACTTCCAGTTGAATTGAATGTAATCACT 4125
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   47 CTGTTGAGAATAATGGAAGCCAGGAAAGGGCAGAGAGAACTTCCAGTTGAATTGAATGTAATCACT 106

Qy   4126 CATGA 4130
           |||||
Db   107 CATGA 111

RESULT 4
LOCUS   BM330353   416 bp   mRNA   linear   EST 04-JAN-2002
DEFINITION PIC1_50_G12_g1_A002 Pathogen-infected compatible 1 (PIC1) Sorghum
            bicolor cDNA, mRNA sequence.
ACCESSION   BM330353
VERSION   BM330353.1   GI:18069490
KEYWORDS   EST.
SOURCE   sorghum.
  ORGANISM Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1 (bases 1 to 416)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Fang,G.C., Dean,R., Wing,R.,
            Sudman,M. and Pratt,L.H.
  TITLE   An EST database from Sorghum: plants infected with a compatible
            pathogen
  JOURNAL   Unpublished (2002)
  COMMENT   Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector, and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20. Three-prime sequences, which are obtained with PolyTMix or
            T7 sequencing primer, are presented as the reverse complement.
            Seq primer: PolyTMix
            High quality sequence start: 484
            POLYA=No.

FEATURES   Location/Qualifiers
     source   1..484
                     /organism="Sorghum bicolor"
                     /db_xref="taxon:4558"

RESULT 5

LOCUS   BI416347   549 bp   mRNA   linear   EST 14-AUG-2001
DEFINITION LjNEST2d3r Lotus japonicus nodule library, mature and immature
            nodules Lotus Japonicus cDNA 5', mRNA sequence.
ACCESSION   BI416347
VERSION   BI416347.1 GI:15186602
KEYWORDS   EST.
SOURCE   Lotus japonicus.
ORGANISM   Lotus japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
            Lotus.
REFERENCE   1 (bases 1 to 549)
AUTHORS   Freund,S., Stougaard,J. and Udvardi,M.
TITLE   Lotus Japonicus root nodule ESTs: a tool for functional genomics
JOURNAL   Unpublished (2000)
COMMENT   Contact: Udvardi MK
            Molecular Plant Nutrition
            Max Planck Institute of Molecular Plant Physiology
            Am Muehlenberg 1, 14476 Golm, Germany
            Fax: 49 331 567 8250
            Email: udvardi@mpimp-golm.mpg.de
            Seq primer: T7
            High quality sequence stop: 549.
FEATURES   Location/Qualifiers
     source   1..549
                     /organism="Lotus japonicus"
                     /cultivar="Gifu (B-129)"
                     /db_xref="taxon:34305"
                     /clone_lib="Lotus Japonicus nodule library, mature and
                     immature nodules"
                     /note="Organ: Nodule; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; The library was prepared using mRNA
                     extracted from a mixture of mature (pink) and immature
                     (white) nodules of Lotus Japonicus ecotype 'Gifu'. Nodules
                     were induced by, and contained Mesorhizobium loti strain
                     NZP2235."
BASE COUNT   178 a   110 c   123 g   138 t
ORIGIN

Qy   4153 GGTCAAAGGCAGCTTGCTGAGAGAGTACTATCGAGTCTCCAAGGATGA 4199
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   86 GGTCAAAGGCAGCTTGCTGAGAGAGTACTATCGAGTCTCCAAGGATGA 132

Qy   4153 GGTCAAAGGCAGCTTGCTGAGAGAGTACTATCGAGTCTCCAAGGATGA 4199

BASE COUNT   121 a   75 c   100 g   120 t
ORIGIN

Qy   4153 Score 47;   DB 13;   Length 416;
Best Local Similarity   100.0%;   Pred. No. 2.1e-11;
Matches   47;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Query Match   1.0%;

/clone_lib="Pathogen-infected compatible 1 (PIC1)"
/tissue_type="Leaves"
/dev_stage="4-week-old seedlings infected with
Colletotrichum graminicola"
/note="Vector: pBluescript II SK(-) from Lambda Zap II;
Site_1: XhoI; Site_2: EcoRI; Four-week-old sorghum
seedlings were sprayed with spore suspension prepared from
3-week-old FRM42I, a sorghum isolate of the anthracnose
pathogen Colletotrichum graminicola. Inoculated plants
were kept in a 25 C dark growth chamber with 100% relative
humidity for 24 hr, followed by 12/12 hr of light/dark
cycle at 25 C with 90% relative humidity for another 24
hr. All leaves were harvested and quick frozen with liquid
nitrogen and stored in a -80 C freezer. The library was
made from poly-A RNA in the cloning vector lambda ZAP II.
Clones to be sequenced were prepared by mass excision.
WARNING: While most or all ESTs are expected to derive
from the host plant, no effort was made to eliminate ESTs
deriving from the pathogen."

RESULT 6

LOCUS   BE443510   458 bp   mRNA   linear   EST 25-JUL-2000
DEFINITION WHE1115_D09_H17zs Wheat etiolated seedling root normalized cDNA
            library Triticum aestivum cDNA clone WHE1115_D09_H17, mRNA
            sequence.
ACCESSION   BE443510
VERSION   BE443510.1 GI:9443044
KEYWORDS   EST.
SOURCE   bread wheat.
ORGANISM   Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1 (bases 1 to 458)
AUTHORS   Anderson,O.D. Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
            ,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Nguyen,H.T.,
            Rausch,C.J., Seaton,C.L., Tong,J.C. and Zhang,D.
TITLE   The structure and function of the expressed portion of the wheat
            genomes - Normalized root cDNA library
JOURNAL   Unpublished (2000)
COMMENT   Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Stratagene SK primer.
FEATURES   Location/Qualifiers
     source   1..458
                     /organism="Triticum aestivum"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="WHE1115_D09_H17"
                     /clone_lib="Wheat etiolated seedling root normalized cDNA
                     library"
                     /tissue_type="Root"
                     /dev_stage="Five day old etiolated seedling"
                     /lab_host="E. coli DH10B"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid
                     pBluescript SK; Site_1: EcoRI; Site_2: XhoI; Seeds were
                     surface-sterilized, germinated and grown aseptically in
                     the dark at room temperature on filter paper with water,
                     nystatin and cefotaxime in covered crystallization
                     dishes. Roots were harvested. The tissue, total RNA, and
                     poly(A) RNA were prepared, a cDNA library was made in the
                     TJ Close lab (Choi, Close, Fenton) at the University of
                     California, Riverside. The cDNA clones were in vivo
                     excised to give pBluescript phagemids before
                     normalization was carried out. The mass excision of the
                     phagemid library and normalization were done in HT Nguyen
                     lab by D. Zhang at Texas Tech University. Normalization
                     protocol used was that of Soares. Plasmid DNA
                     preparations and DNA sequencing were performed in the OD
                     Anderson lab (all other authors)."
BASE COUNT   173 a   88 c   94 g   103 t
ORIGIN

Qy   3662 TCCAGCTTAAGACAACTGAAATGGCAAACAAGGAC 3696
             |||||||||||||||||||||||||||||||||||
Db   157 TCCAGCTTAAGACAACTGAAATGGCAAACAAGGAC 191

Query Match   0.8%;

Best Local Similarity   100.0%;   Pred. No. 1.6e-05;
Matches   35;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   3662 TCCAGCTTAAGACAACTGAAATGGCAAACAAGGAC 3696

Score 35;   DB 13;   Length 549;

Query Match   0.7%;

Best Local Similarity   100.0%;   Pred. No. 0.0047;
Matches   30;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Score 30;   DB 10;   Length 458;

```
Qy   1345  GGAAAGCTCCAGGCAGAAGCTGATGCTCAC 1374
            |||||||||||||||||||||||||||||
Db    291  GGAAAGCTCCAGGCAGAAGCTGATGCTCAC 320
```

Search completed: December 26, 2002, 08:39:20
Job time : 3880 secs

OM protein - protein search, using sw model

Copyright (c) 1993 - 2002 Compugen Ltd.

GenCore version 5.1.3

Run on: December 26, 2002, 08:39:26 ; Search time 82 Seconds
                 (without alignments)
                 3306.805 Million cell updates/sec

Title: US-09-538-396-2
Perfect score: 1316
Sequence: 1 MSTVDKMLIKGIRSFDPDNK................YRVSKDENQHSIIESQEIFD 1316

Scoring table: OLIGO
                 Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 20647115 residues

Word size :   20

Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :   SPTREMBL_21:*
    1: sp_archea:*
    2: sp_bacteria:*
    3: sp_fungi:*
    4: sp_human:*
    5: sp_invertebrate:*
    6: sp_mammal:*
    7: sp_mhc:*
    8: sp_organelle:*
    9: sp_phage:*
   10: sp_plant:*
   11: sp_rodent:*
   12: sp_virus:*
   13: sp_vertebrate:*
   14: sp_unclassified:*
   15: sp_rvirus:*
   16: sp_bacteriap:*
   17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	49	3.7	1316	10	Q9SLO2	Q9SLO2 arabidopsis
2	42	3.2	1312	11	Q9JIL8	Q9JIL8 rattus norv
3	42	3.2	1312	11	P70388	P70388 mus musculu
4	31	2.4	1173	4	Q9UP86	Q9UP86 homo sapien
5	31	2.4	1312	4	Q92878	Q92878 homo sapien
6	31	2.4	1318	4	O43254	O43254 homo sapien

ALIGNMENTS

RESULT 1
ID   Q9SLO2   PRELIMINARY;   PRT;   1316 AA.
AC   Q9SLO2; Q9M6P9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)

DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Putative RAD50 DNA repair protein (DNA repair-recombination
DE   protein).
GN   AT2G31970 OR RAD50.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. COLUMBIA;
RX   MEDLINE=21097002; PubMed=11169180;
RA   Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA   Barnstead M.E., Bowman C.L., Ronning C.M., Benito M.-I.,
RA   Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA   Fraser C.M., Venter J.C.;
RL   Plant J. 25:31-41(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. COLUMBIA;
RA   Gallego M.E., Jeanneau M., Granier F., Bouchez D., Bechtold N.,
RA   White C.I.;
RT   "Disruption of the Arabidopsis RAD50 gene leads to plant sterility and
RT   MMS sensitivity.";
RL   Plant J. 25:31-41(2001).
RA   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC006223; AAD15407.2; -.
DR   EMBL; AF168748; AAF36810.1; -.
DR   InterPro; IPR003439; ABC_transporter.
DR   InterPro; IPR004584; Rad50.
DR   TIGRFAMs; TIGR00606; rad50; 1.
SQ   SEQUENCE   1316 AA; 152814 MW; 89DC4F6BCA39B0E8 CRC64;

Query Match   3.7%; Score 49; DB 10; Length 1316;
Best Local Similarity   100.0%;
                        Matches   49; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   152 LENVIFVHQDESNWPLQDPSTLKKKFDDIFSATRTRKALEVIKKLHKDQ 200
                  |||||||||||||||||||||||||||||||||||||||||||||||
Db   152 LENVIFVHQDESNWPLQDPSTLKKKFDDIFSATRTRKALEVIKKLHKDQ 200

RESULT 2
ID   Q9JIL8   PRELIMINARY;   PRT;   1312 AA.
AC   Q9JIL8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Rad50.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20368653; PubMed=10908350;
RA   Lanson N.A. Jr., Egeland D.B., Royals B.A., Claycomb W.C.;
RT   "The MRE11-NBS1-RAD50 Pathway is perturbed in SV40 large T antigen-
RT   immortalized AT-1 and HL-1 cardiomyocytes.";
RL   Nucleic Acids Res. 28:2882-2892(2000).
DR   EMBL; AF215576; AAF91229.1; -.
DR   InterPro; IPR003439; ABC_transporter.
DR   InterPro; IPR001395; Aldo/Ket_red.
DR   InterPro; IPR004584; Rad50.
DR   TIGRFAMs; TIGR00606; rad50; 1.
DR   PROSITE; PS00063; ALDOKETO_REDUCTASE_3; UNKNOWN_1.
SQ   SEQUENCE   1312 AA; 153783 MW; F13C041BD2C05932 CRC64;

Query Match   3.2%; Score 42; DB 11; Length 1312;
Best Local Similarity   100.0%;
Matches   42; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1207 MRGRCSAGQKVLASLIIRLALAETFCLNCGILALDEPTTNLD 1248
            ||||||||||||||||||||||||||||||||||||||||||
Db   1197 MRGRCSAGQKVLASLIIRLALAETFCLNCGILALDEPTTNLD 1238

RESULT 3
ID   P70388;   PRELIMINARY;   PRT;   1312 AA.
AC   P70388;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   RAD50.
GN   RAD50 OR RAD50.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97067183; PubMed=8910585;
RA   Kim K.K., Daud A.I., Wong S.C., Pajak L., Tsai S.C., Wang H.,
RA   Henzel W.J., Field L.J.;
RT   "Mouse RAD50 has limited epitopic homology to p53 and is expressed in
RT   the adult myocardium.";
RL   J. Biol. Chem. 271:29255-29264(1996).
DR   EMBL; U66887; AAC52894.1; -.
DR   MGD; MGI:109292; Rad50.
DR   InterPro; IPR003439; ABC_transporter.
DR   InterPro; IPR001395; Aldo/ket_red.
DR   InterPro; IPR004584; Rad50.
DR   TIGRFAMs; TIGR00606; rad50; 1.
PROSITE; PS00063; ALDOKETO_REDUCTASE_3; UNKNOWN_1.
SQ   SEQUENCE   1312 AA;   153487 MW;   4AF9AF9AD9E1D7A2 CRC64;

Qy   1207 MRGRCSAGQKVLASLIIRLALAETFCLNCGILALDEPTTNLD 1248
            ||||||||||||||||||||||||||||||||||||||||||
Db   1197 MRGRCSAGQKVLASLIIRLALAETFCLNCGILALDEPTTNLD 1238

Query Match   3.2%;   Score 42;   DB 11;   Length 1312;
Best Local Similarity 100.0%;   Pred. No. 5.3e-32;
Matches   42;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

RESULT 4
ID   Q9UP86;   PRELIMINARY;   PRT;   1173 AA.
AC   Q9UP86;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Truncated RAD50 protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99346140; PubMed=10415333;
RA   Kim K.K., Shin B.A., Seo K.H., Kim P.N., Koh J.T., Kim J.H.,
RA   Park B.R.;
RT   "Molecular cloning and characterization of splice variants of human
RT   RAD50 gene.";
RL   Gene 235:59-67(1999).
DR   EMBL; AF057300; AAD50326.1; -.
DR   InterPro; IPR003439; ABC_transporter.
DR   InterPro; IPR001395; Aldo/ket_red.
DR   InterPro; IPR004584; Rad50.
DR   TIGRFAMs; TIGR00606; rad50; 1.
PROSITE; PS00063; ALDOKETO_REDUCTASE_3; UNKNOWN_1.
SQ   SEQUENCE   1173 AA;   138432 MW;   D6734A4E4D898AAE CRC64;

Qy   1207 MRGRCSAGQKVLASLIIRLALAETFCLNCGILALDEPTTNLD 1248
            ||||||||||||||||||||||||||||||||||||||||||
Db   1197 MRGRCSAGQKVLASLIIRLALAETFCLNCGILALDEPTTNLD 1238

Query Match   2.4%;   Score 31;   DB 4;   Length 1173;
Best Local Similarity 100.0%;   Pred. No. 3.8e-21;
Matches   31;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1207 MRGRCSAGQKVLASLIIRLALAETFCLNCGI 1237
            ||||||||||||||||||||||||||||||
Db   1058 MRGRCSAGQKVLASLIIRLALAETFCLNCGI 1088

RESULT 5
ID   Q92878;   PRELIMINARY;   PRT;   1312 AA.
AC   Q92878;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   RAD50.
GN   RAD50.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dolganov G.M., Maser R.S., Novikov A., Tosto L., Chong S.,
RA   Bressan D.A., Petrini J.H.J.;
RL   Mol. Cell. Biol. 16:0-0(0).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99346140; PubMed=10415333;
RA   Kim K.K., Shin B.A., Seo K.H., Kim P.N., Koh J.T., Kim J.H.,
RA   Park B.R.;
RT   "Molecular cloning and characterization of splice variants of human
RT   RAD50 gene.";
RL   Gene 235:59-67(1999).
DR   EMBL; U63139; AAB07119.1; -.
DR   EMBL; AF057299; AAD50325.1; -.
DR   InterPro; IPR003439; ABC_transporter.
DR   InterPro; IPR001395; Aldo/ket_red.
DR   InterPro; IPR004584; Rad50.
DR   TIGRFAMs; TIGR00606; rad50; 1.
PROSITE; PS00063; ALDOKETO_REDUCTASE_3; UNKNOWN_1.
SQ   SEQUENCE   1312 AA;   153891 MW;   1F208C3817FC41FC CRC64;

Qy   1207 MRGRCSAGQKVLASLIIRLALAETFCLNCGI 1237
            ||||||||||||||||||||||||||||||
Db   1197 MRGRCSAGQKVLASLIIRLALAETFCLNCGI 1227

Query Match   2.4%;   Score 31;   DB 4;   Length 1312;
Best Local Similarity 100.0%;   Pred. No. 4.2e-21;
Matches   31;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

RESULT 6
ID   O43254;   PRELIMINARY;   PRT;   1318 AA.
AC   O43254;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   RAD50 homologue hsRAD50.
GN   RAD50.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RA   Offenberg H.H.;
RL   Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z75311; CAA99729.1; -.

```
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001395; Aldo/ket_red.
DR   InterPro; IPR004584; Rad50.
DR   TIGRFAMs; TIGR00606; rad50; 1.
DR   PROSITE; PS00063; ALDOKETO_REDUCTASE_3; UNKNOWN_1.
SQ   SEQUENCE   1318 AA;  154410 MW;  B4EB9EEBB105822 CRC64;
```

```
        Query Match             2.4%;  Score 31;  DB 4;  Length 1318;
        Best Local Similarity 100.0%;  Pred. No. 4.2e-21;
        Matches   31;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1207 MRGRCSAGQKVLASLIIRLALAETFCLNCGI 1237
            |||||||||||||||||||||||||||||||
Db     1203 MRGRCSAGQKVLASLIIRLALAETFCLNCGI 1233
```

Search completed: December 26, 2002, 08:49:18
Job time : 83 secs

OM protein - protein search, using sw model

Run on: December 26, 2002, 05:28:26 ; Search time 68 Seconds
                                                (without alignments)
                                                2578.791 Million cell updates/sec

```
                    GenCore version 5.1.3
                    Copyright (c) 1993 - 2002 Compugen Ltd.
```

Title: US-09-538-396-2
Perfect score: 1316
Sequence: 1 MSTVDKMLIKGIRSFDPDNK.........YRVSKDENQHSIESQEIFD 1316

Scoring table: OLIGO
                    GapOp 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size :   20 :

Total number of hits satisfying chosen parameters: 3

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :   A_Geneseq_101002:*

```
              1:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
              2:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
              3:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
              4:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
              5:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
              6:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
              7:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
              8:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
              9:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
             10:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
             11:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
             12:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
             13:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
             14:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
             15:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
             16:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
             17:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
             18:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
             19:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
             20:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
             21:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
             22:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
             23:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
                           %
Result                   Query
No.    Score   Match Length   DB    ID            Description
------ -----   ------------   --    --            -----------
1      1316    100.0   1316   22    AAB27248      Maize RAD50.    Zea
2      31      2.4     1312   18    AAW22775      Human RAD50.    Homo
3      31      2.4     1312   19    AAW71295      Human homologue of
```

ALIGNMENTS

RESULT 1
AAB27248

```
ID   AAB27248 standard; Protein; 1316 AA.
XX
AC   AAB27248;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Maize RAD50.
XX
OS   Zea mays.
XX
KW   Maize; RAD50; meiotic recombination; DNA repair; chromosome 4p;
KW   recombination efficiency; transformation efficiency.
XX
PN   WO200068404-A1.
XX
PD   16-NOV-2000.
XX
PF   25-APR-2000; 2000WO-US11086.
XX
PR   05-MAY-1999;  99US-0132575.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Mahajan PB, Shi J;
XX
DR   WPI; 2001-007402/01.
DR   N-PSDB; AAC66915.
XX
PT   Nucleic acids encoding Rad50 polypeptides, useful for transgenic plant
PT   production and for producing antibodies that are useful for modifying
PT   gene expression -
XX
PS   Claim 11; Page 61-62; 76pp; English.
XX
CC   The present sequence is the protein sequence of maize RAD50, the gene for
CC   which is found on maize chromosome 4p. RAD50 is involved in meiotic
CC   recombination and DNA repair, and the protein and its gene can be used to
CC   produce transgenic plants whose expression of the gene may be regulated.
CC   This can be useful in the regulation of transformation and recombination
CC   efficiency in plants.
XX
SQ   Sequence   1316 AA;
```

```
Query Match             100.0%;  Score 1316;  DB 22;  Length 1316;
Best Local Similarity   100.0%;
Matches  1316; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

Qy     1 MSTVDKMLIKGIRSFDPDNKRNVITFFKPLTLIVGPNGAGKTTIIECLKLSCTGELPPNSR 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MSTVDKMLIKGIRSFDPDNKRNVITFFKPLTLIVGPNGAGKTTIIECLKLSCTGELPPNSR 60

Qy    61 SGHTFVHDPKVAGETETKGQIKLRFKTAAGKDVVCIRSFQLTQRASKMEFKAIESVLQTI 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 SGHTFVHDPKVAGETETKGQIKLRFKTAAGKDVVCIRSFQLTQRASKMEFKAIESVLQTI 120

Qy   121 NPHTGERVCLSYRCADMDREIPALMGVSKAVLENVIFVHQDESNMPLQDPSTLKKKFFDI 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 NPHTGERVCLSYRCADMDREIPALMGVSKAVLENVIFVHQDESNMPLQDPSTLKKKFFDI 180

Qy   181 FSATRYTKALEVIKKLHKDQMQEITKFRLKLENLQTVKDQAHKLRENIAQDQEKSDASKS 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 FSATRYTKALEVIKKLHKDQMQEITKFRLKLENLQTVKDQAHKLRENIAQDQEKSDASKS 240

Qy   241 QMEQLKEKICGTERELLQMETSLDELRRLGQGIDIKATERSTLLTQQHERKLAALSEENED 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 QMEQLKEKICGTERELLQMETSLDELRRLGQGIDIKATERSTLLTQQHERKLAALSEENED 300

Qy   301 TDEELMEWQTKFEERIALLLEFKISKLVRDMDDEASYSSVLSKQNSELTHEIGKLQAEADA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TDEELMEWQTKFEERIALLLEFKISKLVRDMDDEASYSSVLSKQNSELTHEIGKLQAEADA 360

Qy   361 HLTMKHERDSDIKNICTKHNLGPVPEHPFPTNDVAMNLTNRIKARLSSLENDLLDKKKSNE 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   361  HLTMKHERDSDIKNICTKHNLGPVPEHPFTNDVAMNLTNRIKARLSSLENDLLDKKKSNE  420

Qy   421  DQLDVLWKHYLKINARYSEVDGQIQSKIESNSGILRRRKDKEKERDAAEVELSKFNLSRI  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  DQLDVLWKHYLKINARYSEVDGQIQSKIESNSGILRRRKDKEKERDAAEVELSKFNLSRI  480

Qy   481  DERERINQIEVERKTLALGERDIDSIISQKRTEVISLEQKIKVLLREKDIINRNADERVK  540
          ||||||||||||||||||||||||||||||| ||||| |||||||||||||||||||||
Db   481  DERERINQIEVERKTLALGERDYDSIISQKRFEVVSLEQKIKVLLREKDIINRNADERVK  540

Qy   541  LGLKDALESSKDKLNEIVNEHKDKIKKVLRGRNPFEKDMKKEINQAFWPVDKEYNELRS   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  LGLKDALESSKDKLNEIVNEHKDKIKKVLRGRNPFEKDMKKEINQAFWPVDKEYNELRS   600

Qy   601  KSQEAEQELKFTQSKVTDAREQLTKLRRDMDAKRRFLDSKLQSILQISANVDMFPKVLQD  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  KSQEAEQELKFTQSKVTDAREQLTKLRRDMDAKRRFLDSKLQSILQISANVDMFPKVLQD  660

Qy   661  AMNKRDEQKRLENFANGMREMLAPFEHLARKNHVCPCCERAFTPDEEDEFVKKQRMQNSS  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  AMNKRDEQKRLENFANGMREMLAPFEHLARKNHVCPCCERAFTPDEEDEFVKKQRMQNSS  720

Qy   721  TAERSKALAMESSNAEALFQQLDKLRIYIDAVKLVEETIPLAEKNLNQHLADESQKAQA   780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  TAERSKALAMESSNAEALFQQLDKLRIYIDAVKLVEETIPLAEKNLNQHLADESQKAQA   780

Qy   781  FDDLLGVLAHVQMDRDAVEALLQPTDIIDRHVHEIQQLVKEVEDLEYALDSSGRGVKSLE  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  FDDLLGVLAHVQMDRDAVEALLQPTDIIDRHVHEIQQLVKEVEDLEYALDSSGRGVKSLE  840

Qy   841  EIQLELNFLQRTRDTLIVEVDDLRDQHRMLNEDMSSAQVRWHNAREEKVKASSILERFQK  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  EIQLELNFLQRTRDTLIVEVDDLRDQHRMLNEDMSSAQVRWHNAREEKVKASSILERFQK  900

Qy   901  SEEELVLLAEEKEQLIVERKLLEESLDPLSKERESTLQEYNALKQKLDEEYHQLAERRE   960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  SEEELVLLAEEKEQLIVERKLLEESLDPLSKERESTLQEYNALKQKLDEEYHQLAERRE   960

Qy   961  FQQELDALGRLNMKIKGYLDSKRNEKLKELQGRHVLCHSQLQSCMAKQQRISAELNKSE  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  FQQELDALGRLNMKIKGYLDSKRNEKLKELQGRHVLCHSQLQSCMAKQQRISAELNKSE  1020

Qy  1021  LLQSQGQLKRNIDDNLKRKTKADVEQLTRDIESLEERLLSIGSLSAIEADLKRHSQEKE  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  LLQSQGQLKRNIDDNLKRKTKADVEQLTRDIESLEERLLSIGSLSAIEADLKRHSQEKE  1080

Qy  1081  RLNSEFNRWQGTLSVYQSNISKHHQELKLSQYKDIEKRYTNQFLQLKTTEMANKDLDRY  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  RLNSEFNRWQGTLSVYQSNISKHHQELKLSQYKDIEKRYTNQFLQLKTTEMANKDLDRY  1140

Qy  1141  TALDKALMRFHSMKMEEINKIIKELWQQTYRGQDIDYISINSDSEGAGTRSYSYRVVMQT  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  TALDKALMRFHSMKMEEINKIIKELWQQTYRGQDIDYISINSDSEGAGTRSYSYRVVMQT  1200

Qy  1201  GDAELEMRGRCSAGQKVLASLIIRLALAETFCLNCGILALDEPTTNLDGPNAESLAAALL  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  GDAELEMRGRCSAGQKVLASLIIRLALAETFCLNCGILALDEPTTNLDGPNAESLAAALL  1260

Qy  1261  RIMEARKGQENFQLIVITHDERFAHLIGQRQLAEKYYRVYSKDENQHSIIESQEIFD  1316
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  RIMEARKGQENFQLIVITHDERFAHLIGQRQLAEKYYRVYSKDENQHSIIESQEIFD  1316
```

```
Query Match            2.4%;   Score 31;   DB 18;   Length 1312;
Best Local Similarity  100.0%;  Pred. No. 5.7e-21;
Matches   31;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy  1207  MRGRCSAGQKVLASLIIRLALAETFCLNCGI  1237
          |||||||||||||||||||||||||||||||
Db  1197  MRGRCSAGQKVLASLIIRLALAETFCLNCGI  1227
```

```
PI   Dolganov G;
XX
DR   WPI; 1998-481207/41.
DR   N-PSDB; AAV59979.
XX
PT   Novel human immunomodulatory poly:peptide(s) - have homology to the
PT   yeast RAD50 or Drosophila Septin-2 proteins
XX
PS   Disclosure; Pages 136-140; 155pp; English.
XX
CC   The present sequence represents a human homologue of the yeast
CC   S. cerevisiae gene RAD50. The present sequence has 35% overall
CC   homology to the yeast RAD50 gene, and is expressed in activated
CC   T-cells, testis, foetal liver and heart tissues. The specification
CC   also describes sequences encoding human homologues of the
CC   yeast RAD50, the Drosophila Septin-2 and Acyl-CoA synthetase. The
CC   proteins have immunomodulatory activity. The nucleic acids and
CC   proteins can be used to identify activated T-cells in a sample
CC   population. They can also be used to isolate and identify sequences
CC   encoding other proteins or other compounds having immunomodulatory
CC   activity.
XX
SQ   Sequence   1312 AA;

Query Match             2.4%;    Score 31;   DB 19;   Length 1312;
Best Local Similarity   100.0%;  Pred. No. 5.7e-21;
Matches    31;  Conservative    0;  Mismatches    0;  Indels   0;  Gaps   0;

Qy   1207  MRGRCSAGQKVLASLIIRLALAETFCLNGCI  1237
           |||||||||||||||||||||||||||||||
Db   1197  MRGRCSAGQKVLASLIIRLALAETFCLNGCI  1227
```

Search completed: December 26, 2002, 08:45:17
Job time : 70 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

Run on: December 26, 2002, 05:07:47 ; Search time 7537 Seconds
                 (without alignments)
                 17345.071 Million cell updates/sec

Title: US-09-538-396-1
Perfect score: 4492
Sequence: 1 aattcggcacgagtggatcc..........atctagcatcaaaaaaaaa 4492

Scoring table: OLIGO_NUC
                 Gapop 60:0 , Gapext 60.0

Searched: 2054640 seqs, 1455102878 residues

Word size :   30

Total number of hits satisfying chosen parameters: 3

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :   GenEmbl:*
        1: gb_ba:*
        2: gb_htg:*
        3: gb_in:*
        4: gb_om:*
        5: gb_ov:*
        6: gb_pat:*
        7: gb_ph:*
        8: gb_pl:*
        9: gb_pr:*
       10: gb_ro:*
       11: gb_sts:*
       12: gb_sy:*
       13: gb_un:*
       14: gb_vi:*
       15: gb_ba:*
       16: em_fun:*
       17: em_hum:*
       18: em_in:*
       19: em_mu:*
       20: em_om:*
       21: em_or:*
       22: em_ov:*
       23: em_ph:*
       24: em_pat:*
       25: em_pl:*
       26: em_ro:*
       27: em_sts:*
       28: em_un:*
       29: em_vi:*
       30: em_vrt:*
       31: em_htg_hum:*
       32: em_htg_inv:*
       33: em_htg_other:*
       34: em_htg_mus:*
       35: em_htg_pln:*
       36: em_htg_rod:*
       37: em_htg_mam:*
       38: em_htg_vrt:*
       39: em_sy:*
       40: em_htgo_mus:*
       41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB	ID	Description
1	4492	100.0	4492	6	AX047235	AX047235 Sequence
c	2	38	0.8	124610	2	AP004086	AP004086 Oryza sat
3	38	0.8	136073	2	AP004773	AP004773 Oryza sat

ALIGNMENTS

RESULT 1
LOCUS   AX047235   4492 bp   DNA   linear   PAT 15-DEC-2000
DEFINITION   Sequence 1 from Patent WO0068404.
ACCESSION   AX047235
VERSION   AX047235.1   GI:11876516
KEYWORDS
SOURCE   Zea mays.
  ORGANISM   Zea mays
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
             clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 4492)
  AUTHORS   Mahajan,P.B. and Shi,J.
  TITLE   Maize rad50 orthologue and uses of therof
  JOURNAL   Patent: WO 0068404-A 1 16-NOV-2000;
             PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
     source   1..4492
                     /organism="zea mays"
                     /db_xref="taxon:4577"
     CDS   292..4242
                     /codon_start=1
                     /note="unnamed protein product"
                     /protein_id="CAC18962.1"
                     /db_xref="GI:11876517"
                     /translation="MSTVDKMLIRGIRSFDPDNKNVITFFKPLTLIVGPNGAGKTII
                     ECLKLSCTGELPPNSRSGHTEVHDPKVAGEIPTKGIKLRFKTAAGKDVVCIRSFQLI
                     QKASKMFKAIESVLQINPHPGEKVCLSYRCADMDREIPALMGVSKAVLENVIFVHQ
                     DESNWPLQDPSTLRKKFPDDIFSATRTKALEVIKLHKDQMOEIHFRLKLENLQTVK
                     DQAHLRENIAQDQEKSDASKSQMEQLKEKICGTEREILQMETSLDELRRLOGQIDIK
                     ATERSTLLTQQHEKLAALSEENEDTDELMEWQTKFEERIALLETKISKLVRDMDDEA
                     SYSSVLSKQNSELTHIIGKLQAEADAHILTMKHERDSDIKNICTKHNLGPYEHPPTND
                     VAMLTNRIKARLSLENDLLDKKKSNEDQLDVLWRHYLKINARVSEVDGIQSKIES
                     MSGILRRKDKREKERDAAEVELSKFNLSRIDRERNMQIEVERKTLALGERDYDSIIS
                     OKTEVYSLEQKIKIVLLREKDIINRNADERVKLGLKKDALESSKQKLNEIVNEHKDKI
                     KKVLRGRNPFEKDMKKEINQAFWPVDKEVNELRSKSGQEAEGQELKFTQSKVTDARQGLT
                     KLRRDMDAKRRPLDSKLQSILQISANVDMFPKVLQDAMNKRDEQKRLENFANGMREML
                     OKRTEYYSLEQKIKIVLLREKDIINRNADERVKLGLKKDALESSKQKLNEIVNEHKDKI
                     APFEHLARKNHVCPCCERAFFPDEDEFVKKORMONSSTAERSKALAMESSNAEALFQ
                     OLDKLRTIYDAYVKLVEETIPLAEKNIQHILADESQKAQAPDDLLGVLAHVQMRDAV
                     EALLQPDTIDRHVHEIOLVEVEDLEYALDSSGRGVKSLEEIOLENFLQRTRDTLL
                     IVEVDLLRDQHRMLNEDMSSAQVRWHNARSERVKASSILRFROKSEEELVLLAEKEQ
                     NKIRGILDSKINEKLKRELQGRVLCHSOLQSCMARQORISAELNKSKELLQGOGOLKR
                     LIVEKKLIEESDPLSKEKESLIQEYNALRQKLDEETHQLAERKRFQOELDALGRLN
                     NIDDNLKYRKRADYEQLTTDIESLEEERLLSIGSLSAIEADLKRHSOERLNSEFNR
                     WQGTLSVYQOSNISKHRQELKLSOYKDIEKRVTNOFLQLKTTEMANRDLDRYTALDKA
                     LMRFHSMKEEINKIIKELWQOTYRGQDIDYISINSDSEGAGTRSYSIRVVMQOTGDAE
                     LEMRRCRSAGQKVLASLIIIRLALAEFFCLNGGILALDEPTTNLDGPNAESLAAAALLRI
                     MEARKQOENFQLIVITHDERPAHLIGQRQLAEKYIRVSKDENQHSIIESQEIFD"
BASE COUNT   1537 a   831 c   1061 g   1063 t
ORIGIN
Qy   1 AATTCGGCACGAGTGGATCCATTAGCACCCATAGCCGTACAAAACCCTAAGAAACCCTAAC 60

Query Match   100.0%; Score 4492; DB 6; Length 4492;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4492; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 AATTCGGCACGAGTGGATCCATTAGCACCCATAGCCGTACAAAACCCTAAGAAACCCTAAC 60

```
Db      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1 AATTCGGCACGAGGTGGATCAATTAGACACCTACAAAAACCCTAAC              60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CGGTACAAAAACCCTACCGCGCTGAAGACTCCAAAAAAAACGCGAATTTTCTC      120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   61 CGGTACAAAAACCCTACCGCGCTGAAGACTCCAAAAAAAACGCGAATTTTCTC      120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CTCCACTGCCCCTCCTTTTCTCTTTCCAATCGTTTTGCAATCACTACGAGCTTAATGAAT  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  121 CTCCACTGCCCCTCCTTTTCTCTTTCCAATCGTTTTGCAATCACTACGAGCTTAATGAAT  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AGAAGTTGATAGGGAGAGATAGCATCCGCAAATCTAGGTTTGGGCAATCGGCTCTGGCCAGAC  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  181 AGAAGTTGATAGGGAGAGATAGCATCCGCAAATCTAGGTTTGGGCAATCGGCTCTGGCCAGAC  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 TGGATCGGAGTGCAAGTCGTAGAGGAGGCACTGAAGATGAGCACC              300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  241 TGGATCGGAGTGCAAGTCGTAGAGGAGGCACTGAAGATGAGCACC              300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 GTTGACAAGATGCTGATCAAAGGGGATTCGGGAGCTTCAGACACTCATC          360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  301 GTTGACAAGATGCTGATCAAAGGGGATTCGGGAGCTTCAGACACTCATC          360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 ACCTTCTTCAAGCCGCTCACCCTCATCGTTGGGCCCAACGGTGCTGGGGAAGACCACGATC  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  361 ACCTTCTTCAAGCCGCTCACCCTCATCGTTGGGCCCAACGGTGCTGGGGAAGACCACGATC  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 CGGTTTAAGACTGCAGCAGGAAAGGATGTGGTGTGTCATCGGGCCTCCAGCTTACCCAA    480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  421 CGGTTTAAGACTGCAGCAGGAAAGGATGTGGTGTGTCATCGGGCCTCCAGCTTACCCAA    480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 ACCTTCGTCCACGACCCCAAAGGTAGCTGGCGGAAACAAAATTAAGTTG          540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  481 ACCTTCGTCCACGACCCCAAAGGTAGCTGGCGGAAACAAAATTAAGTTG          540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 ATCGAGTGCCTGAAGCTTTCTTGCACCGGGGAGCTGCCCAACTCCCGGCTCTGGCCAC    600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  541 ATCGAGTGCCTGAAGCTTTCTTGCACCGGGGAGCTGCCCAACTCCCGGCTCTGGCCAC    600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 AAGGCCATCAAAGATGGAGTTTAAGGCAATTGAAGCGTCCTCCCAGACTATAAATCCACAC  660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  601 AAGGCCATCAAAGATGGAGTTTAAGGCAATTGAAGCGTCCTCCCAGACTATAAATCCACAC  660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 ACAGGGGAGAAAAGTCTGCCCTCAGCGTACAGATGTGCTGACAGTGTGGATAGAGAAGAGTTCCTGCC  720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  661 ACAGGGGAGAAAAGTCTGCCCTCAGCGTACAGATGTGCTGACAGTGTGGATAGAGAAGAGTTCCTGCC  720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 TTAATGGGTGGTTTCGAAGGCCGTACTGGAGAGGTGTCATCCCGGTTCCTCCAAGATGAATCC     780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  721 TTAATGGGTGGTTTCGAAGGCCGTACTGGAGAGGTGTCATCCCGGTTCCTCCAAGATGAATCC     780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 AATTGGCCATTGCAGGACCCGTCAACACTTAAGCAGAAGTTCGATGCACACTCTTCTGTCGCC    840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  781 AATTGGCCATTGCAGGACCCGTCAACACTTAAGCAGAAGTTCGATGCACACTCTTCTGTCGCC    840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 ACACGCTATACGGAAAAGCTCTTGAAGTCATAAAGAAAACTTCACAGAGGATCAAATGCAAGAG       900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  841 ACACGCTATACGGAAAAGCTCTTGAAGTCATAAAGAAAACTTCACAGAGGATCAAATGCAAGAG       900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 ATCAAGACTTTTAGGTTAAAGCTGGAGAACCTTCAGCACTGTAAAAAGACCAAGCACATAAG        960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  901 ATCAAGACTTTTAGGTTAAAGCTGGAGAACCTTCAGCACTGTAAAAAGACCAAGCACATAAG        960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 CTGCGTGAAAATATTGCTCAAGATCAAGAAAAGTCAGATGCCTCAAAATCTCAGATGGAG        1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  961 CTGCGTGAAAATATTGCTCAAGATCAAGAAAAGTCAGATGCCTCAAAATCTCAGATGGAG        1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 CAACTGAAAGGAAAAGACTCTGTGGGTACCGAGAGAAATCCTGCAAATGGAAACAAGTTTG     1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1021 CAACTGAAAGGAAAAGACTCTGTGGGTACCGAGAGAAATCCTGCAAATGGAAACAAGTTTG     1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 GATGAACTGAGGAGGACTTCAGGGACAAAATTGACATCAAGGCAACAGAGAAGTACATTA     1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1081 GATGAACTGAGGAGGACTTCAGGGACAAAATTGACATCAAGGCAACAGAGAAGTACATTA     1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 1081 GATGAACTGAGGAGGACTTCAGGGACAAAATTGACATCAAGGCAACAGAGAAGTACATTA     1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1141 CTTACGCAGCAGCATGAGAAAGCTTGCCTGACACTTTCTGAGGAAAAATGAAGAGTACGATGAG 1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 CTTACGCAGCAGCATGAGAAAGCTTGCCTGACACTTTCTGAGGAAAAATGAAGAGTACGATGAG 1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1201 GAACTAAGGGAAAGGGCAAACAAAAATTTGAAGAAAGGAGATTGCGTTAACTGAGGAAACAAAAATC 1260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 GAACTAAGGGAAAGGGCAAACAAAAATTTGAAGAAAGGAGATTGCGTTAACTGAGGAAACAAAAATC 1260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1261 AGTAAACTTGTTAAGAGAATATGGATGAAGAAGCAACTTATAGTCCGTTCTTCCAAACAA   1320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 AGTAAACTTGTTAAGAGAATATGGATGAAGAAGCAACTTATAGTCCGTTCTTCCAAACAA   1320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1321 AATTCTGAATTAACACATGAAAATTGGAAAGCTCCAGGCAGCAGAAGCTCACCTGACT     1380
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 AATTCTGAATTAACACATGAAAATTGGAAAGCTCCAGGCAGCAGAAGCTCACCTGACT     1380
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1381 ATGAAGCACGAAACGAGACGTCCAGACAACACATAAAATATATGCACTGAACAGAAAAATCAATAATCTTGGGCCG 1440
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 ATGAAGCACGAAACGAGACGTCCAGACAACACATAAAATATATGCACTGAACAGAAAAATCAATAATCTTGGGCCG 1440
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1441 GTTCCTGAACATCCCTTTACGAAATGATGTTGCTAGGAAACCTTACAAAACAGGATTAAAGCG    1500
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441 GTTCCTGAACATCCCTTTACGAAATGATGTTGCTAGGAAACCTTACAAAACAGGATTAAAGCG    1500
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1501 AGACTATCAAGTCTTGAAGATGAATTTGCTGGATAAGAAGAAAATCCAAATGAAGATCAGTTA    1560
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501 AGACTATCAAGTCTTGAAGATGAATTTGCTGGATAAGAAGAAAATCCAAATGAAGATCAGTTA    1560
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1561 ATGAAGCACGAAACGAGACCTCAGACACTAAACATAAAATGTCTAACTAAACATAATCGCTACTGCAGAAGTTGATGGTCAG 1620
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561 ATGAAGCACGAAACGAGACCTCAGACACTAAACATAAAATGTCTAACTAAACATAATCGCTACTGCAGAAGTTGATGGTCAG 1620
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1621 ATACAAATCTAAGATTGAATTCCATGTCAACATTTTAAGACCGGAGAAAAAGATAAAGAGAAA    1680
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621 ATACAAATCTAAGATTGAATTCCATGTCAACATTTTAAGACCGGAGAAAAAGATAAAGAGAAA    1680
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1681 GAACGCGGATGCTGCGAGAAGTGGGAGCTTTCAAAATTTAAATCTATCCCGGTATCGTGGATGAGAGG    1740
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681 GAACGCGGATGCTGCGAGAAGTGGGAGCTTTCAAAATTTAAATCTATCCCGGTATCGTGGATGAGAGG    1740
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1741 GAGAGACATATGCAATTGAAGTCGAGAGAAGGAAGACACTTGCGCTTGGAGTGAAGGGACTAT   1800
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741 GAGAGACATATGCAATTGAAGTCGAGAGAAGGAAGACACTTGCGCTTGGAGTGAAGGGACTAT   1800
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1801 GATTCAATTATAAGTCAGAAACGAACAGAGAAGTGTATATAGTTGGAACAGAAAATAAAAGTG    1860
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801 GATTCAATTATAAGTCAGAAACGAACAGAGAAGTGTATATAGTTGGAACAGAAAATAAAAGTG    1860
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1861 CTTCTGCGGGAGAAAGATATAAATGAGAGTAAAATGGGTTTG    1920
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861 CTTCTGCGGGAGAAAGATATAAATGAGAGTAAAATGGGTTTG    1920
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1921 AAGAAGGAGGATGCAATTGGAAAGAGCAGCAAGGACAAGGACAAATCCTTTGACTTAATGAGCATTAAG 1980
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1921 AAGAAGGAGGATGCAATTGGAAAGAGCAGCAAGGACAAGGACAAATCCTTTGACTTAATGAGCATTAAG 1980
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1981 GATAAAATCAAAAAGGTTACTTAGGGGAGGAATCCTTTTGAGAAGGATTATGATGAAGAGGAG    2040
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1981 GATAAAATCAAAAAGGTTACTTAGGGGAGGAATCCTTTTGAGAAGGATTATGATGAAGAGGAG    2040
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 2041 ATCAATCAAGACCTTTTGGCCTGTGGGACAAGGAATAACAATGACTTAAACAATCAAAATCCCAG 2100
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2041 ATCAATCAAGACCTTTTGGCCTGTGGGACAAGGAATAACAATGACTTAAACAATCAAAATCCCAG 2100
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 2101 GAAGCAGAGCAGAGAGCCTTAAATTTACTCAGACAAGGAGTAAACTGAATGCATTAATTG   2160
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2101 GAAGCAGAGCAGAGAGCCTTAAATTTACTCAGACAAGGAGTAAACTGAATGCATTAATTG   2160
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 2161 ACAAAACTTCGAAGAGATATGGAGGAATGGATGCAAAAGAGAGATTCCTGGACTCGAAAACTTCAATCT 2220
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2161 ACAAAACTTCGAAGAGATATGGAGGAATGGATGCAAAAGAGAGATTCCTGGACTCGAAAACTTCAATCT 2220
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   2221 ATTTTACAGATATCTGCTAATCTGAATGTTGCATGTTGTTTCCCAAAGTTCTACAAGACGCCATGAAC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 ATTTTACAGATAATCTGCTAATCTGAATGTTTGCATGTTGTTTCCCAAAGTTCTACAAGACGCCATGAAC 2280

Qy   2281 AAAAGAGATGAACAGAAAAGATTAGAGAATTTCGCAAATGCCTTCCA 2340
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 AAAAGAGATGAACAGAAAAGATTAGAGAATTTCGCAAATGCCTTCCA 2340

Qy   2341 CCTTTTGAACATTTGGCTCGGAAGAATCATGTTATGCCCATGCTTGCTTTCACA 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 CCTTTTGAACATTTGGCTCGGAAGAATCATGTTATGCCCATGCTTGCTTTCACA 2400

Qy   2401 CCTGATGAGGAGGGATGAGTTCGTGAAGAAAACTCAAGTACTGCAGAG 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 CCTGATGAGGAGGGATGAGTTCGTGAAGAAAACTCAAGTACTGCAGAG 2460

Qy   2461 AGATCTAAAGCTCTGGCAATGGAATCATCAAATGCTGAAGCTGCTTTTTTCAGCAATTGGAT 2520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 AGATCTAAAGCTCTGGCAATGGAATCATCAAATGCTGAAGCTGCTTTTTTCAGCAATTGGAT 2520

Qy   2521 AAAACTTCGGACTATCTATGATGATGCTTATGTGAAGCTGGTAGAAGAAAACCATACCTCTAGCA 2580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 AAAACTTCGGACTATCTATGATGATGCTGGTTATGTGAAGCTGGTAGAAGAAAACCATACCTCTAGCA 2580

Qy   2581 GAGAAAAACTTGAATCAACATTGGGGGGATGAAAGTCAGAAGGCGCAGGCATTTGATGAT 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 GAGAAAAACTTGAATCAACATTGGGGGGATGAAAGTCAGAAGGCGCAGGCATTTGATGAT 2640

Qy   2641 CTTTTGGGTGTTCTTGCCCATGTTCAAAATGCTGGCAGGGACGTCAGTGGAAGCCTTATTACAA 2700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 CTTTTGGGTGTTCTTGCCCATGTTCAAAATGCTGGCAGGGACGTCAGTGGAAGCCTTATTACAA 2700

Qy   2701 CCCCACTACTAATTGACAGGCCATGTACATGAAAATTCAACAGCTAGTACAGAAGAAGTAGAA 2760
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 CCCACTGATACTATTGACAGGCCATGTGTGAAAGCTGGTCAAAGAAGAAGTAGAA 2760

Qy   2761 GATCTTGAATATGCACTTGATTCTAGTGCGGTGTCAAGTCTTGGAGGAAAATTCAA 2820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 GATCTTGAATATGCACTTGATTCTAGTGCGGTGTCAAGTCTTGGAGGAAAATTCAA 2820

Qy   2821 CTGGAGCTGAACTTTCTGCGAGAACACAGGACACATTGATTGTGAAGGATGGATGATCTT 2880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 CTGGAGCTGAACTTTCTGCGAGAACACAGGACACATTGATTGTGAAGGATGGATGATCTT 2880

Qy   2881 AGAGATCAACATAGAAATGCTAAATGAAGATAATGTCAAGTGCTCAGGTGAGATGGCACAAT 2940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 AGAGATCAACATAGAAATGCTAAATGAAGATAATGTCAAGTGCTCAGGTGAGATGGCACAAT 2940

Qy   2941 GCTCGGGAGAGAAGTGAAAGCTTCTACCATTCCAAAAATCTCAAGAG 3000
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 GCTCGGGAGAGAAAGTGAAAGCTTCTACCATTCCAAAAATCTGAAAG 3000

Qy   3001 GAATTGGTGGTCTTCTTAGCTGTGAGAGAAAAAGAGCTTTTAGAA 3060
          |||||||||||||||||||||||||||||||||||||||||||||
Db   3001 GAATTGGTGGTCTTCTTAGCTGTGAGAGAAAAAGAGCTTTTAGAA 3060

Qy   3061 GAGTCTCTTGATCCATTGTCCAAAGAGAAAAGAGAGCTTGTTGCAAGAGTATAATGCTTTG 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 GAGTCTCTTGATCCATTGTCCAAAGAGAAAAGAGAGCTTGTTGCAAGAGTATAATGCTTTG 3120

Qy   3121 AAGCAAAAGCTGGATGATATCATCAGCTTTGCAGAGAAAAAAGGGAGTTCCAGCAA 3180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 AAGCAAAAGCTGGATGAGAGTATCATCAGCTTTGCAGAAGAAAAAGGGAGTTCCAGCAA 3180

Qy   3181 GAACTTGATGCTCTTGGAAGACTTAATATGAAGATAAAAAGGGTTACTTGGATTCCAAGAAA 3240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 GAACTTGATGCTCTTGGAAGACTTAATATGAAGATAAAAAGGGTTACTTGGATTCCAAGAAA 3240

Qy   3241 AACGAAAAGCTTAAGGAATTGCAGGAGGAGGCATGTTCTTGCTTCTTTGCCATTCTCAGTTACAGAGT 3300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241 AACGAAAAGCTTAAGGAATTGCAGGAGGAGGCATGTTCTTGCTTCTTTGCCATTCTCAGTTACAGAGT 3300
```

```
Qy   3301 TGCATGGCAAAAACAGCAGCAAAGAGAATATCAGAGTTGAAGTTAAAACAGCAGCTACTGCAG 3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301 TGCATGGCAAAAACAGCAGCAAAGAGAATATCAGAGTTGAAGTTAAAACAGCAGCTACTGCAG 3360

Qy   3361 GGCCAGGGCCAGTTGAAAAGACAGCAAGAAAGCAAGTACAGAGAGAAAAAAGGCT 3420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 GGCCAGGGCCAGTTGAAAAGACAGCAAGAAAGCAAGTACAGAGAGAAAAAAGGCT 3420

Qy   3421 GATGTGGGAACAACTTACTCGTGATGATAGGAAGATACTTGAAGGAAAAAGGT 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 GATGTGGGAACAACTTACTCGTGATGATAGGAAGATACTTGAAGGAAAAAGGT 3480

Qy   3481 AGCTTGTCTGCCTATAGAAGCTGATCGTGAAACGCCAACTCTCCAAGAGAATCACTTAAT 3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481 AGCTTGTCTGCCTATAGAAGCTGATCGTGAAACGCCAACTCTCCAAGAGAATCACTTAAT 3540

Qy   3541 TCAGAATTTAACAGGTGGCAAGAACACTTTCTGTTTATCAAAAGTAAATATTCAAAGCAC 3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541 TCAGAATTTAACAGGTGGCAAGAACACTTTCTGTTTATCAAAAGTAAATATTCAAAGCAC 3600

Qy   3601 AACAAGAGCTTAAACTGTCACAGTATACAAGGATAATCGAGAAGGCCGATATACTAATCAATTT 3660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3601 AACAAGAGCTTAAACTGTCACAGTATACAAGGATAATCGAGAAGGCCGATATACTAATCAATTT 3660

Qy   3661 CTCCAGCTTAAGACAACTGAAAATGGCAAAACAGGTGGACAGTAGTTACTGCTTTA 3720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3661 CTCCAGCTTAAGACAACTGAAAATGGCAAAACAGGTGGACAGTAGTTACTGCTTTA 3720

Qy   3721 TCAGAATTTATGCGGTTCCACAGCATGTGCAAGGAACACTTTCTGTTTATCAAAAGTAAATATAATC 3780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721 GACAAGGCTCTTATGCGGTTCCACAGCATGTGCAAGGAACACTTTCTGTTTATCAAAAGTAAATATAATC 3780

Qy   3781 GAACTGTGGCAACAGACAATACAGAGGGCCAGGATATTGAATTACATAGAGCGATAAATTCTGAT 3840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 GAACTGTGGCAACAGACAATACAGAGGACCAGGATATTGAATTACATAGAGCGATAAATTCTGAT 3840

Qy   3841 TCTGAGGGGTGCCCACTCGATCAATCAGATCAGGGCCCAGGGCTTTGGCTTTGTGAGAATGGCA 3900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841 TCTGAGGGGTGCTGGGCACTCGATCAATCAGATCAGGGCCCAGGGCTTTGGCTTTGTGAGAATGG 3900

Qy   3901 GAGCTGGAAATCGCGAGGGCGCGCTGCAGTGCTCAGAAGGTTCTGCGAAAGGTTCGAAGGTTCGAAG 3960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901 GAGCTGGAAATCGCGAGGGCGCGCTGCAGTGCTCAGAAGGTTCTGCGAAAGGTTCGAAGGTTCGAAG 3960

Qy   3961 AGACTAGCACTTGCGAGAAACTTTCTGCCTGACCTGCGGGTATATTGGCTTGGCTTGGAGCCA 4020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961 GAACTGTGGCAACAGACAATACAGAGGACCAGGATATTGGCTTGGAGATGGCCA 4020

Qy   4021 ACTACGAAATCTAGATGGGCCCAAATGCAGAGTGTTCTGCTCTGCTGTTGAGAATGTAATG 4080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 ACTACGAATCTAGATGGGCCCAAATGCAGAGTGTTCTGCTCTGCTGTTGAGAATGTAATG 4080

Qy   4081 GAAGCTGGAAAATGCGAGGGCGCGCTGCAGTGCTCAGAAGGTTCTGCTTCTCTTATAATC 4140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 GAAGCTGGAAATGCGAGGGCGCGCTGCAGTGCTCAGAAGGTTCTGCTTCTCTTATAATC 4140

Qy   4141 GCCCATCTTATCGGTCAAAAGGCAGCAAGCTTGCTGAGAAGTACTACTCCAAGGATGAG 4200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 GCCCATCTTATCGGTCAAAAGGCAGCAAGCTTGCTGAGAAGTACTACTCCAAGGATGAG 4200

Qy   4201 AACCACAGCACAGCTATAATTGAATCCCAAGAGAATATTGACAAAGGGTCGTTCTAGGAGGCTG 4260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 AACCACAGCACAATTATCGGTCAAAAGGCAGCAAGCTTGCTGAGAAGTACTACTCCAAGGATGAG 4260

Qy   4261 TAGCAGGCACACTCGTTTTGCTAGTCGAATCCAGTAGTAATTATGCCAAGTACTACTGGGTGCCAGAG 4320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261 TAGCAGCACACTCGTTTTGCTAGTCGAATCCAGTAGTAATTATGCCAAGTACTACTGGGTGCCAGAG 4320

Qy   4321 CAATGTTAACAAGCTTTAGGAGGCTCTGTTACGTCAGTTGCTGAAACCTAT 4380
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4321 CAATGTTAACAAGCTTTAGGAGGCTCTGTTACGGCGTTACGTCAGTTGCTGAAACCTAT 4380

Qy   4381 CCTTCGTTGTTGTATACTTAATTTAATCTGGCCACAGGATAGGAATGTGTGCACTGGGTGAT 4440
```

```
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4381 CCTTCGTTGTTGTTATTACTTATTAATCTGGCCAGAGGATGGAATGTGTGCACTGGGTGAT 4440

Qy    4441 GGATGTTTCACGACATCAATGAATGTTTCACAATCTAGCATCAAAAAAAAAA 4492
           ||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    4441 GGATGTTTCACGACATCAATGAATGTTTCACAATCTAGCATCAAAAAAAAAA 4492
```

RESULT 2

Db	61062

LOCUS AP004086/c

DEFINITION Oryza sativa (japonica cultivar-group)
 OJ1288_D09, *** SEQUENCING IN PROGRESS ***, in ordered pieces.

ACCESSION AP004086

VERSION AP004086.1 GI:15281365

KEYWORDS HTG; HTGS_PHASE2.

SOURCE Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
 clone:OJ1288_D09.

ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1 (bases 1 to 124610)

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

TITLE Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC
 clone:OJ1288_D09

JOURNAL Published Only in Database (2001)

REFERENCE 2 (bases 1 to 124610)

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

TITLE Direct Submission

JOURNAL Submitted (22-AUG-2001) Takuji Sasaki, National Institute of
 Agrobiological Resources, Rice Genome Research Program; Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
 Tel:81-298-38-7441, Fax:81-298-38-7468)

COMMENT The nucleotide sequence of this BAC clone was generated by
 combining Monsanto and RGP-Japan sequencing data.
 NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is
 available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES Location/Qualifiers

source 1..124610
 /organism="Oryza sativa (japonica cultivar-group)"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="2"
 /clone="OJ1288_D09"

BASE COUNT 35712 a 27346 c 25887 g 35511 t 154 others

ORIGIN

Qy 3892 GGTGATGCTGAGCTGGAAATGCGAGGCGCGCTGCAGTGC 3929
   |||||||||||||||||||||||||||||||||||||||
Db 61062 GGTGATGCTGAGCTGGAAATGCGAGGCGCGCTGCAGTGC 61025

RESULT 3

LOCUS AP004773

DEFINITION Oryza sativa (japonica cultivar-group) chromosome 2 clone
 *** SEQUENCING IN PROGRESS ***, in ordered pieces.

ACCESSION AP004773

VERSION AP004773.1 GI:18844993

Query Match 0.8%; Score 38; DB 2; Length 124610;
Best Local Similarity 100.0%; Pred. No. 4.4e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

---

```
KEYWORDS HTG; HTGS_PHASE2.

SOURCE Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
 clone:P0416F07.

ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

TITLE Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, PAC
 clone:P0416F07

JOURNAL Published Only in Database (2002)

REFERENCE 2  (bases 1 to 136073)

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

TITLE Direct Submission

JOURNAL Submitted (20-FEB-2002) Takuji Sasaki, National Institute of
 Agrobiological Sciences, Rice Genome Research Program; Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
 Tel:81-298-38-7441, Fax:81-298-38-7468)

COMMENT NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is
 available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES Location/Qualifiers

source 1..136073
 /organism="Oryza sativa (japonica cultivar-group)"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="2"
 /clone="P0416F07"

BASE COUNT 38865 a 28336 c 29834 g 38561 t 477 others

ORIGIN

Qy  3892 GGTGATGCTGAGCTGGAAATGCGAGGCGCGCTGCAGTGC 3929
   |||||||||||||||||||||||||||||||||||||||
Db 108120 GGTGATGCTGAGCTGGAAATGCGAGGCGCGCTGCAGTGC 108157

Query Match 0.8%; Score 38; DB 2; Length 136073;
Best Local Similarity 100.0%; Pred. No. 4.4e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Search completed: December 26, 2002, 07:46:52
Job time : 8293 secs

LOCUS 136073 bp DNA linear HTG 21-FEB-2002
LOCUS 124610 bp DNA linear HTG 21-MAR-2002